

Stochastic differential games for a multiclass M/M/1 queueing problem with model uncertainty

Asaf Cohen*

February 22, 2017

Abstract

We consider a multidimensional stochastic differential game that emerges from a multiclass M/M/1 queueing problem under heavy-traffic with model uncertainty. Namely, it is assumed that the decision maker is uncertain about the rates of arrivals to the system and the rates of service and acts to optimize an overall cost that accounts for this uncertainty. We show that the multidimensional game can be reduced to a one-dimensional stochastic differential game. Then, the value function of the games is characterized as the unique solution to a free-boundary problem from which we extract equilibria for both the reduced and the multidimensional games. Finally, we analyze the dependence of the value function and the equilibria on the ambiguity parameters.

AMS Classification: Primary: 93E20, 60K25, 91A15, 60J60; secondary: 49J15, 35R35.

Keywords: multiclass M/M/1, Brownian control problem, stochastic differential game, ambiguity aversion, model uncertainty, heavy-traffic, the Harrison–Taksar free-boundary problem.

1 Introduction

We consider a multiclass M/M/1 queueing model under diffusion-scaled heavy-traffic where the decision maker (DM) is uncertain about the parameters and acts to optimize an overall cost that accounts for this uncertainty. The model consists of a server that at any time instant its effort is allocated by the DM to costumers from several number of classes. Customers of each class are kept in a finite buffer. Apart from the scheduling control, upon arrival of a costumer, the DM has to decide whether to reject it or to assign it to the buffer that corresponds to its class type. The DM has ambiguity about the rates of arrivals and the mean service times. The cost accounts for the ambiguity, the holding of customers in the buffers, and rejections of new arrivals.

The problem without ambiguity was analyzed by [6], under the framework of G/G/1. Plambeck et al. studied in [25] a similar non-robust problem with time constraints instead of the finite buffers constraints. In these problems as well as in many other classical models of queueing control problems, see e.g., [8, 11, 12] and the references therein, there is a fixed random

*Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA. shloshim@gmail.com
web: <https://sites.google.com/site/asafcohentau/>

model; that is, the DM is certain about the evolution of the system, which, moreover does not change in time. Such an assumption is not realistic, and a robust analysis is desirable. Recently, queueing systems under the moderate-deviation heavy-traffic regime with risk sensitive performance criteria were studied, see e.g., [1, 9, 2, 5, 3, 4]. This setup models a ‘very’ robust DM that considers rare events. In fact, the model from [6] mentioned earlier, was studied in the moderate-deviation heavy-traffic regime in [2, 3]

We assume that based on the available data, the DM has a reference model in mind, which, up to some degree, describes the situation he is facing. To model the uncertainty about the reference model, the DM takes into account other models and penalizes them based on their deviation from the reference model. The penalization depends then on how averse the DM is to ambiguity. Such ambiguity models are sometimes referred to as *model uncertainty* or *Knightian uncertainty*, see e.g., [24, 14, 7] and in the context of queueing systems see [21, 10, 23]. We allow for different levels of model uncertainty for each of the arrival processes as well as for each of the service processes. More specifically, we consider the following cost function, which the DM aims to minimize.

$$\sup_{\mathbb{Q}} \left\{ \mathbb{E}^{\mathbb{Q}} \left[\int_0^{\infty} e^{-\rho t} \left(\hat{h} \cdot X(t) dt + \hat{r} \cdot dR(t) \right) \right] - \sum_{i=1}^I \frac{1}{\kappa_{i,1}} L^{\rho}(\mathbb{Q}_{i,1} \| \mathbb{P}_{i,1}) - \sum_{i=1}^I \frac{1}{\kappa_{i,2}} L^{\rho}(\mathbb{Q}_{i,2} \| \mathbb{P}_{i,2}) \right\},$$

where I is the number of classes. The vectors \hat{h} and \hat{r} stand for the holding and the rejection costs, respectively. The I -dimensional processes X and R represent the queue lengths and the rejection process, respectively. The measures $\mathbb{P}_{i,1}$ and $\mathbb{P}_{i,2}$ are the reference probability measure of the arrival and the departure processes. The parameters $\kappa_{i,j} > 0$ are the ambiguity parameters, which penalize the deviation from the reference measures to the measure $\mathbb{Q}^n = \prod_{i=1}^I (\mathbb{Q}_{i,1} \times \mathbb{Q}_{i,2})$. The function L^{ρ} is a variant of the Kullback–Liebler divergence.

Typically, heavy-traffic queueing problems in the diffusion scale are treated by defining a limiting control problem associated with Brownian motion, called *Brownian control problems*, first introduced by [15]; for further reading on Brownian control problems see e.g., [8, 11, 12] and the references therein. In our case, the cost function given above suggests that the limiting problem should be in fact a stochastic game. The players in this game are the DM and the nature, which according to their goals are referred to as the *minimizer* and the *maximizer*, respectively. Interpreting the roles of the processes from the queueing control problem (QCP) to a *multidimensional stochastic differential game* (MSDG), the minimizer controls the server’s effort allocation and the admission/rejection, while the maximizer may perturb the drift of the Brownian motion (possibly, differently for each coordinate), which is derived from the arrival and service rates. The game’s cost consists of the original holding and rejection penalties and a variant of the Kullback–Liebler divergence with respect to (w.r.t.) the relevant measures in this setup. The latter component stands for a penalty for the maximizer for changing the drift. At this point it is worth mentioning that in the moderate-deviation heavy-traffic regime models mentioned above, the limit behavior is governed by a deterministic differential game with players that have the same roles as in our stochastic game. Moreover, in both of these games and ours, the maximizer is penalized for changing the drift; see the dynamics in (3.7) and the cost function in (3.12) here vs. [2, Equation (11)] and the cost function given below that.

The paper is devoted to deriving and studying the stochastic game, while the convergence of the value of the QCP to the value of the MSDG is the subject of a work in progress [13].

We show that a *state-space collapse* property holds. That is, we provide a one-dimensional game, called the *reduced stochastic differential game* (RSDG), which its state is the workload process. The roles of the two players remain the same and the dynamics and the cost functions have similar components. We show that the games are equivalent in the sense that for any strategy of the minimizer in either one of the games, we construct a strategy for the minimizer in the other game that performs at least as well, and therefore, also the value functions are the same. For further reading about workload reduction, see [15, 19, 16, 17, 20]. The advantage of such a reduction is that the dynamics live in a lower dimension and have only two components of singular controls. Therefore, most of the analysis is performed in the RSDG setup. We characterize the value function of the RSDG as the unique classical solution of a *Hamilton-Jacoby-Bellman* (HJB) equation. By this, we extend the relationship between the reduced Brownian control problem and a relevant HJB equation studied in [18, Equation (1.2)] and in [6, Equation (41)] to a similar relationship in a stochastic game setup with a different HJB. Unlike in [6], due to the existence of a maximal player in our model, the HJB is not linear. Therefore, one cannot use the results given in these papers and rather needs to establish the relationship between the two. As a first step, we reduce the problem of solving the HJB equation to a free-boundary problem. Then we use the *shooting method* to solve the latter problem. In short, this is a method that is used to solving boundary value problems by a reduction to initial value problems and we take it one step forward in the free-boundary setup; see [29, Section 7.3] for further reading about the method. Then we show that the optimal strategy for the minimizer in this game is a reflecting strategy. Meaning, that the minimizer should use minimal idleness and minimal amount of rejections in order to keep the workload in a fixed interval. From this policy we construct an optimal policy for the minimizer in the MSDG, according to which, the queue length processes evolve along a certain curve in the state space. Moreover, we supply equilibria in both games. The equilibria strategies for the minimizers in our games are shown to have the same structures as the optimal strategies in the risk-neutral Brownian control problems in [6]. Such a result is not so obvious for reasons having to do with the non-stationarity structure of the problem caused by the existence of the maximizer player. Moreover, we show that the equilibria in our case share similarities with the equilibria of the deterministic differential games from [2]. The difference between the policy of the minimizer in the multidimensional setup here as compared to [6] and [2] emerges from the cutoff point of the reflecting strategy in the RSDG, which affects the point of reflection on the multidimensional curve. This suggests that the policy, which has a very simple structure, is good for various levels of ambiguity, starting from the non-robust problem, going through our formulation of robustness here, to the ‘very’ robust moderate-deviation heavy-traffic regime. The DM only has to decide, what is his ambiguity level and choose the reflecting point in accordance.

Aside from studying the games, we also analyze the dependence of the games on the ambiguity parameters. We show continuity of the value function and the optimal reflecting strategies w.r.t. the ambiguity parameters and that as the ambiguity vanishes, the problem converges to the risk-neutral problem studied in [6].

In summary, our main contributions are as follows. We

- introduce a new type of Brownian control problem, which is in fact a stochastic game that governs the limiting behavior of a multiclass M/M/1 queueing problem with model uncertainty;

- show that a state-space collapse property holds for this game;
- show that the reduced game solves uniquely a relevant HJB equation, which is a nonlinear free boundary problem and that there is an optimal reflecting strategy (Theorem 4.1);
- provide equilibria for the two games considered (Theorem 5.1);
- analyze the dependence of the value function and the equilibria on the ambiguity parameters (Theorems 6.1 and 6.2).

The paper is organized as follows. In Section 2 we present the queueing model, the scaling regime, and the cost criteria. In Section 3 we motivate and present the stochastic differential games and study the relationship between the two. Next, In Section 4 we study the WBCP. We provide the HJB equation and prove that the value function of the game is the unique smooth solution of the HJB. Moreover, we show that the minimizer has an optimal reflecting strategy. In Section 5 we discuss about uniqueness of the optimal reflecting strategy and find equilibria in both games. Finally, in Section 6 we study the dependency on the ambiguity parameters.

1.1 Notation

We use the following notation. For $a, b \in \mathbb{R}$, $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. For a positive integer k and $c, d \in \mathbb{R}^k$, $c \cdot d$ denotes the usual scalar product and $\|c\| = (c \cdot c)^{1/2}$. We denote $[0, \infty)$ by \mathbb{R}_+ . For subintervals $I_1, I_2 \subseteq \mathbb{R}$ and $m \in \{1, 2\}$ we denote by $\mathcal{C}(I_1, I_2)$, $\mathcal{C}^m(I_1, I_2)$, and $\mathcal{D}(I_1, I_2)$ the space of continuous functions [resp., functions with continuous derivatives of order m , functions that are right-continuous with finite left limits (RCLL)] mapping $I_1 \rightarrow I_2$. The space $\mathcal{D}(I_1, I_2)$ is endowed with the usual Skorohod topology.

2 The queueing model

2.1 The reference probability space and some fundamental processes

The model consists of I customer classes and a single server. Each class has its own finite buffer and upon arrival, customers are queued in the corresponding buffer or rejected. Within each class, customers are served at the order of their arrivals. Processor sharing is allowed and the server may serve up to I customers at a time, where two customers from the same class cannot be served simultaneously. The system is studied under heavy-traffic. For this, we consider a sequence of systems, indexed by the scaling parameter $n \in \mathbb{N}$. For every $i \in \mathcal{I} := \{1, \dots, I\}$ and $n \in \mathbb{N}$ we consider a *reference probability space* $(\Omega_{i,1}^n, \mathcal{G}_{i,1}^n, \mathbb{P}_{i,1}^n)$ that supports a Poisson process A_i^n with a given rate λ_i^n . It counts the number of arrivals to the i -th buffer.

In a similar way, consider a probability space $(\Omega_{i,2}^n, \mathcal{G}_{i,2}^n, \mathbb{P}_{i,2}^n)$ that supports a Poisson process S_i^n with rate μ_i^n . This is the *potential service time* process. That is, for every $t \in \mathbb{R}_+$, $S_i^n(t)$ is the number of service completions of class i customers had server i worked for t units of time.

The fundamental reference probability space that supports the processes $A^n = (A_i^n)_{i=1}^I$ and $S^n = (S_i^n)_{i=1}^I$ is given by,

$$(\Omega^n, \mathcal{G}^n, \mathbb{P}^n) := \left(\prod_{i=1}^I (\Omega_{i,1}^n \times \Omega_{i,2}^n), \otimes_{i=1}^I (\mathcal{G}_{i,1}^n \otimes \mathcal{G}_{i,2}^n), \prod_{i=1}^I (\mathbb{P}_{i,1}^n \times \mathbb{P}_{i,2}^n) \right),$$

where $\otimes_{i=1}^I (\mathcal{G}_{i,1}^n \otimes \mathcal{G}_{i,2}^n) = (\mathcal{G}_{1,1}^n \otimes \mathcal{G}_{1,2}^n) \otimes \dots \otimes (\mathcal{G}_{I,1}^n \otimes \mathcal{G}_{I,2}^n)$. The reason that we consider such a decomposed probability space is because we allow for different levels of ambiguity for the arrival and service rates in the system. This issue will become clearer in Section 2.2

Notice that from the structure of the probability space it follows that for every fixed $n \in \mathbb{N}$, under the measure \mathbb{P}^n , the processes $A_1^n, S_1^n, \dots, A_I^n, S_I^n$ are mutually independent. Moreover, the distribution of A_i^n (resp., S_i^n) under \mathbb{P}^n is identical to its distribution under $\mathbb{P}_{i,1}^n$ (resp., $\mathbb{P}_{i,2}^n$).

Let $U^n = (U_i^n)_{i=1}^I$ be an RCLL process taking values in $\mathbb{U} = \{x = (x_1, \dots, x_I) \in [0, 1]^I : \sum x_i \leq 1\}$, where $U_i^n(t)$ represents the fraction of effort devoted at time t by the server to the class- i customer at the head of the line. For each $i \in \mathcal{I}$, the process

$$T_i^n(t) := \int_0^t U_i^n(s) ds, \quad t \in \mathbb{R}_+, \quad (2.1)$$

represents the units of time that the server devoted to class i until time t . For every $t \in \mathbb{R}_+$ and $i \in \mathcal{I}$, $S_i^n(T_i^n(t))$ is the number of service completions of class i customers until time t . This is a Cox process with intensity $\mu_i^n U_i^n$. Rejections of customers are allowed upon arrival and a rejected customer will never return to the system. The number of customers from class i that were rejected by time t is denoted by $R_i^n(t)$ and satisfies,

$$R_i^n(t) = \int_0^t z_i^n(s) dA_i^n(s), \quad t \in \mathbb{R}_+,$$

for some process z_i^n . For every $i \in \mathcal{I}$, the balance equation is given by,

$$X_i^n(t) = X_i^n(0) + A_i^n(t) - S_i^n(T_i^n(t)) - R_i^n(t), \quad t \in \mathbb{R}_+, \quad (2.2)$$

where $X_i^n(t)$ stands for the number of class i customers in the system at time t . For simplicity, we assume that $X_i^n(0)$ is deterministic. We use the notation $X^n = (X_i^n)_{i=1}^I$, and similarly for R^n and T^n . Recall that U^n is an RCLL process, and by construction, A^n and S^n are also RCLL processes. Therefore, so are X^n and R^n .

We assume that

$$\lambda_i^n := \lambda_i n + \hat{\lambda}_i n^{1/2} + o(n^{1/2}), \quad \mu_i^n := \mu_i n + \hat{\mu}_i n^{1/2} + o(n^{1/2}), \quad (2.3)$$

where $\lambda_i, \mu_i \in (0, \infty)$ and $\hat{\lambda}_i, \hat{\mu}_i \in \mathbb{R}$ are fixed. Moreover, the system is assumed to be *critically loaded*, that is, $\sum_{i=1}^I \rho_i = 1$, where $\rho_i := \lambda_i / \mu_i$, $i \in \mathcal{I}$.

The scaled version of (2.2) is given by,

$$\hat{X}_i^n(t) = \hat{X}_i^n(0) + \hat{m}_i^n t + \hat{A}_i^n(t) - \hat{S}_i^n(T_i^n(t)) + \hat{Y}_i^n(t) - \hat{R}_i^n(t), \quad t \in \mathbb{R}_+,$$

where

$$\begin{aligned} \hat{X}^n(t) &:= n^{-1/2} X^n(t), \quad \hat{A}^n(t) := n^{-1/2} (A^n(t) - \lambda^n t), \quad \hat{S}^n(t) := n^{-1/2} (S^n(t) - \mu^n t), \\ \hat{Y}_i^n(t) &:= \mu_i^n n^{-1/2} (\rho_i t - T_i^n(t)), \quad \hat{R}^n(t) := n^{-1/2} R^n(t), \end{aligned} \quad (2.4)$$

and

$$\hat{m}_i^n := n^{-1/2} (\lambda_i^n - \rho_i \mu_i^n).$$

As previously, we use the notation $\hat{Y}^n = (\hat{Y}_i^n)_{i=1}^I$ and $\hat{m}^n = (\hat{m}_i^n)_{i=1}^I$.

The capacity of buffer i is given by $b_i^n := b_i n^{1/2}$ for some constant $b_i \in (0, \infty)$, $i \in \mathcal{I}$. We assume that $X_0^n \in \mathcal{X}$, and the rejection mechanism assures that

$$\hat{X}_i^n(t) \in \mathcal{X}, \quad t \in \mathbb{R}_+, \quad \mathbb{P}^n\text{-a.s.}, \quad (2.5)$$

where $\mathcal{X} := \prod_{i=1}^I [0, b_i]$. It will be assumed throughout that, for some $x_0 \in \mathcal{X}$,

$$\lim_{n \rightarrow \infty} \hat{X}^n(0) \rightarrow x_0. \quad (2.6)$$

The process (U^n, R^n) is regarded as a control in the n -th system and is now given rigorously.

Definition 2.1 (admissible control for the decision maker, QCP) *An admissible control for the minimizer for any initial state $\hat{X}^n(0)$ is a process (U^n, R^n) taking values in $\mathbb{U} \times \mathbb{R}_+^I$ that satisfies the following,*

(i) *(U^n, R^n) is adapted to the filtration $\mathcal{G}_t^n = \mathcal{G}^n(t) := \sigma\{A_i^n(s), S_i^n(T_i^n(s)), i \in \mathcal{I}, s \leq t\}$ and has RCLL sample paths;*

(ii) *the process R^n is nondecreasing in each of its coordinates;*

(iii) *for each $i \in \mathcal{I}$ and $t \geq 0$,*

$$X_i^n(t) = 0 \quad \text{implies} \quad U_i^n(t) = 0;$$

(iv) *the buffer constraint (2.5) holds.*

The first condition expresses the fact that the DM makes his decision based on past observations. The second condition follows since rejections are accumulated. The third condition asserts that servers cannot be active when their buffer is empty. We denote the set of admissible controls for the DM in the n -th system by $\hat{\mathcal{A}}^n(\hat{X}^n(0))$.

2.2 The robust optimization problem

We now turn to describing the cost function. Fix $\varrho > 0$, $\hat{h} \in (0, \infty)^I$, $\hat{r} \in (0, \infty)^I$. The *risk-neutral optimization problem* studied in [6] is given by

$$\inf_{(U^n, R^n) \in \hat{\mathcal{A}}^n(\hat{X}^n(0))} \mathbb{E}^{\mathbb{P}^n} \left[\int_0^\infty e^{-\varrho t} [\hat{h} \cdot \hat{X}^n(t) dt + \hat{r} \cdot d\hat{R}^n(t)] \right].$$

The vectors \hat{h} and \hat{r} stand for the holding and rejection costs, respectively.

We on the other hand study a robust variant of this problem, where the DM is uncertain about the underlying reference probability measure \mathbb{P}^n , or in other words, he suspects that the rates $\{\lambda_i^n\}_{i=1}^I$ and $\{\mu_i^n\}_{i=1}^I$ may be unspecified or may even change over the time. Therefore, instead of optimizing under the reference measure \mathbb{P}^n , he considers a set of candidate measures

(provided in the sequel) and penalizes their deviation from \mathbb{P}^n . The penalization is done by using a discounted variant of the Kullback–Leibler divergence, given by

$$L^\varrho(\hat{\mathbb{Q}}_{i,j}^n \|\mathbb{P}_{i,j}^n) := \mathbb{E}^{\hat{\mathbb{Q}}_{i,j}^n} \left[\int_0^\infty \varrho e^{-\varrho t} \log \frac{d\hat{\mathbb{Q}}_{i,j}^n(t)}{d\mathbb{P}_{i,j}^n(t)} dt \right], \quad n \in \mathbb{N}, i \in \mathcal{I}, j \in \{1, 2\}. \quad (2.7)$$

To establish the level of ambiguity, for every $i \in \mathcal{I}$, we consider the (finite and) positive parameters $\kappa_{i,1}$ and $\kappa_{i,2}$ that quantify the amount of ambiguity that the DM has regarding the rates λ_i^n and μ_i^n , or in other words, the measures $\mathbb{P}_{i,1}^n$ and $\mathbb{P}_{i,2}^n$, respectively. Set $\kappa := (\kappa_{i,1}, \kappa_{i,2})_{i=1}^I$. The DM is facing the following robust optimization problem:

$$\hat{V}^n(\hat{X}^n(0); \kappa) := \inf_{(U^n, R^n) \in \hat{\mathcal{A}}^n(\hat{X}^n(0))} \sup_{\hat{\mathbb{Q}}^n \in \hat{\mathcal{Q}}^n(\hat{X}^n(0))} \hat{J}^n(\hat{X}^n(0), U^n, R^n, \hat{\mathbb{Q}}^n; \kappa),$$

where

$$\hat{J}^n(\hat{X}^n(0), U^n, R^n, \hat{\mathbb{Q}}^n; \kappa) := \quad (2.8)$$

$$\mathbb{E}^{\hat{\mathbb{Q}}^n} \left[\int_0^\infty e^{-\varrho t} \left(\hat{h} \cdot \hat{X}^n(t) dt + \hat{r} \cdot d\hat{R}^n(t) \right) \right] - \sum_{i=1}^I \frac{1}{\kappa_{i,1}} L^\varrho(\hat{\mathbb{Q}}_{i,1}^n \|\mathbb{P}_{i,1}^n) - \sum_{i=1}^I \frac{1}{\kappa_{i,2}} L^\varrho(\hat{\mathbb{Q}}_{i,2}^n \|\mathbb{P}_{i,2}^n),$$

$\hat{\mathbb{Q}}^n = \prod_{i=1}^I (\hat{\mathbb{Q}}_{i,1}^n \times \hat{\mathbb{Q}}_{i,2}^n)$, and the set of candidate measures $\hat{\mathcal{Q}}^n(\hat{X}^n(0))$ is provided rigorously after the next paragraph.

When $\kappa_{i,j}$ is ‘small’ (resp., ‘big’) we say that there is a weak (resp., strong) ambiguity about the rates of the processes A_i^n and $S_i^n(T_i^n) := S_i^n(T_i^n(\cdot))$. The idea is that for small $\kappa_{i,j}$ ’s there is a big punishment for unit of deviation from the reference measure and therefore, the measures $\hat{\mathbb{Q}}_{i,j}^n$ and $\mathbb{P}_{i,j}^n$ should be close to each other and as a consequence also the relevant expectations. However, one needs to make sure that the total punishment given by $\frac{1}{\kappa_{i,j}} L^\varrho(\hat{\mathbb{Q}}_{i,j}^n \|\mathbb{P}_{i,j}^n)$ is also small. In Section 6 we show that as the ambiguity parameters converge to zero, the stochastic differential games, which are provided in Section 3, converge to the risk-neutral Brownian control problem studied in [6]. Therefore, our problem indeed models ambiguity w.r.t. the risk-neutral model. As mentioned in the introduction, the convergence of the QCP to the MSDG is out of the scope of this paper and is studied in [13]. There we also show that, when $\sup_{i,j} \{\kappa_{i,j}\} \rightarrow 0+$, the queueing control problem converges to the risk-neutral problem from [6]. When $\sup_{i,j} \{\kappa_{i,j}\} \rightarrow \infty$, which stands for a very risk averse DM, the value function goes to infinity.

We now turn to define the set of candidate measures. A probability measure $\hat{\mathbb{Q}}^n = \prod_{i=1}^I (\hat{\mathbb{Q}}_{i,1}^n \times \hat{\mathbb{Q}}_{i,2}^n)$ belongs to $\hat{\mathcal{Q}}^n(\hat{X}^n(0))$ if for every $i \in \mathcal{I}$ and $t \in \mathbb{R}_+$ it satisfies

$$\frac{d\hat{\mathbb{Q}}_{i,1}^n(t)}{d\mathbb{P}_{i,1}^n(t)} = \exp \left\{ \int_0^t \log \left(\frac{\psi_{i,1}^n(s)}{\lambda_i^n(s)} \right) dA_i^n(s) - \int_0^t (\psi_{i,1}^n(s) - \lambda_i^n(s)) ds \right\}, \quad (2.9)$$

$$\frac{d\hat{\mathbb{Q}}_{i,2}^n(t)}{d\mathbb{P}_{i,2}^n(t)} = \exp \left\{ \int_0^t \log \left(\frac{\psi_{i,2}^n(s)}{\mu_i^n(s)} \right) dS_i^n(T_i^n(s)) - \int_0^t (\psi_{i,2}^n(s) - \mu_i^n(s)) dT_i^n(s) \right\}, \quad (2.10)$$

for a $\{\mathcal{G}_t^n\}$ -predictable measurable processes $\psi_{i,j}^n$, $j \in \{1, 2\}$, satisfying for every $t \in \mathbb{R}_+$,

$$\psi_{i,1}^n(t) = \lambda_i^n + \hat{\psi}_{i,1}^n(t)(\lambda_i^n)^{1/2} + o(n^{1/2}), \quad (2.11)$$

$$\psi_{i,2}^n(t) = \mu_i^n + \hat{\psi}_{i,2}^n(t)(\mu_i^n)^{1/2} + o(n^{1/2}), \quad (2.12)$$

with

$$\lim_{n \rightarrow \infty} \left\| \hat{\psi}_{i,j}^n \right\|_{\infty} < \infty, \quad \mathbb{P}^n\text{-a.s.}$$

These conditions assure, first, that the right-hand side (r.h.s.) of (2.9)–(2.10) are $\mathbb{P}_{i,j}^n$ -martingales, $j \in \{1, 2\}$. Second, that under the measure $\hat{\mathbb{Q}}_{i,1}^n$ (resp., $\hat{\mathbb{Q}}_{i,2}^n$), A_i^n (resp., $S_i^n(T_i^n)$) is a counting process with intensity $\psi_{i,1}^n$ (resp., $\psi_{i,2}^n U_i^n$), and third, that under the measures $\hat{\mathbb{Q}}_{i,j}^n$, $j \in \{1, 2\}$, the critically load condition is preserved. As for the latter part, notice that for every $i \in \mathcal{I}$ and $t \in \mathbb{R}_+$,

$$\begin{aligned} \hat{A}_i^n(t) &= n^{-1/2} \left(A_i^n(t) - \int_0^t \psi_{i,1}^n(s) ds \right) + \lambda_i^{1/2} \int_0^t \hat{\psi}_{i,1}^n(s) ds, \\ \hat{S}_i^n(T_i^n(t)) &= n^{-1/2} \left(S_i^n(T_i^n(t)) - \int_0^t \psi_{i,2}^n(s) dT_i^n(s) \right) + \mu_i^{1/2} \int_0^t \hat{\psi}_{i,2}^n(s) dT_i^n(s). \end{aligned} \quad (2.13)$$

In both lines above, under $\hat{\mathbb{Q}}_i^n$, the first term is approximately a standard Brownian motion and the second term approximate a drift (in case it converges). The main reason we included the terms λ_i and μ_i in (2.11) and (2.12) is in order to improve the appearance of the limiting problems that are given in Section 3. It does not impair the generality of the model.

The intuition behind the structure of the optimization problem is as follows. The DM, also referred to as the *minimizer*, chooses a control based on the past observations. He minimizes a cost that takes into account a possible deviation from the reference model. For this, we consider an adverse ‘player’, also referred to as the *maximizer*, who has access to the policy chosen by the minimizer and to the filtration \mathcal{G}_t^n . This player is penalized for deviating from the reference model.

Remark 2.1 (i) Notice that given any $\{\mathcal{G}_t^n\}$ -predictable process $\psi_{i,j}^n$ that satisfies (2.11), the r.h.s. of (2.9)–(2.10) are martingales, and therefore, there exist probability measures $\hat{\mathbb{Q}}_{i,j}^n$, $j \in \{1, 2\}$, such that $\hat{\mathbb{Q}}_{i,j}^n|_{\mathcal{G}_t^n}$ satisfies (2.9) for all $t \in \mathbb{R}_+$, see [30, Lemma 4.2].

(ii) Notice that the QCP is a stochastic game that models a type of worst case scenario. The minimizer chooses a strategy and the maximizer, which is penalized for deviating from the reference measure, responds to this strategy by choosing a worst case scenario. For further reading about the structure of the information in control problems with model uncertainty, the reader is referred to [28].

3 Two stochastic differential games

We approximate the QCP by robust stochastic differential games that are driven by Brownian motions. In Sections 3.1 and 3.2 we present a multidimensional and a one-dimensional stochastic differential games. Then in Section 3.3 we show that the two games share the same value and that given any strategy (for the minimizer) in one of the games, one can construct a strategy in the second game that performs just as well. In Sections 5 and 6 we give more results about the relationship between the games.

3.1 The multidimensional stochastic differential game (MSDG)

The sequence of the scaled and centered $2I$ -dimensional Poisson processes (\hat{A}^n, \hat{S}^n) weakly converges to a $2I$ -dimensional Brownian motion starting at zero, with zero mean and the covariance matrix $\text{Diag}(\lambda_1^{1/2}, \dots, \lambda_I^{1/2}, \mu_1^{1/2}, \dots, \mu_I^{1/2})$. Formally speaking, if the process \hat{Y}^n is of order one as $n \rightarrow \infty$, which is rigorously proven in [13], we get from its definition in (2.4) that $T^n \rightarrow (\rho_1, \dots, \rho_I)$ and therefore, under \mathbb{P}^n , $(\hat{A}_i^n - \hat{S}_i^n(T_i^n))_{i=1}^I$ weakly converges to an I -dimensional Brownian motion starting at zero, with zero mean and the covariance matrix

$$\hat{\sigma} = (\hat{\sigma}_{ij}) := \text{Diag}\left((2\lambda_1)^{1/2}, \dots, (2\lambda_I)^{1/2}\right),$$

where $\hat{S}_i^n(T_i^n) := \hat{S}_i^n(T_i^n(\cdot))$, $i \in \mathcal{I}$.

Recall that in the QCP, an admissible control was of the form (U^n, R^n) . Notice that $(\hat{Y}^n(t), \hat{R}^n(t))$ is uniquely determined by $(U^n(s), R^n(s))_{0 \leq s \leq t}$. In the MSDG we consider two I -dimensional processes, \hat{R} and \hat{Y} that play the roles of instantaneous controls, which stand for the scaled rejection process \hat{R}^n and the scaled idle time process \hat{Y}^n , respectively. Moreover, the ambiguity about the true underlying probability measure in the QCP, which is formulated by a penalty for deviating from the reference measure is translated to the limiting problem as well. In the definition below we refer to two players by their roles as a minimizer and a maximizer even though the roles can only be derived from the cost function, which is presented afterwards.

Definition 3.1 (admissible controls, MSDG) *An admissible control for the minimizer for any initial state $\hat{x}_0 \in \mathcal{X}$ is a filtered probability space*

$$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}) := \left(\prod_{i=1}^I \Omega^i, \mathcal{F}^1 \otimes \dots \otimes \mathcal{F}^I, \{\mathcal{F}_t\}, \prod_{i=1}^I \mathbb{P}_i \right),$$

that supports a process (\hat{Y}, \hat{R}) taking values in $(\mathbb{R}_+^I)^2$ with RCLL sample paths adapted to the filtration $\{\mathcal{F}_t\}$, where $(\Omega^i, \mathcal{F}^i, \{\mathcal{F}_t^i\}, \mathbb{P}_i)$ supports a one-dimensional standard Brownian motion \hat{B}_i adapted to the filtration $\{\mathcal{F}_t^i\}$, $i \in \mathcal{I}$. Moreover, assume that the following properties hold:

(i)

$$\text{for every } i \in \mathcal{I} \text{ and } 0 \leq s < t, \hat{B}_i(t) - \hat{B}_i(s) \text{ is independent of } \mathcal{F}_s^i \text{ under } \mathbb{P}_i; \quad (3.1)$$

(ii)

$$\theta \cdot \hat{Y} \text{ and } \hat{R}_i, i \in \mathcal{I} \text{ are nonnegative and nondecreasing, where } \theta := (\mu_1^{-1}, \dots, \mu_I^{-1}); \quad (3.2)$$

(iii)

$$\hat{X}(t) = \hat{x}_0 + \hat{m}t + \hat{\sigma}\hat{B}(t) + \hat{Y}(t) - \hat{R}(t), \quad t \in \mathbb{R}_+, \quad (3.3)$$

such that

$$\hat{X}(t) \in \mathcal{X}, \quad t \in \mathbb{R}_+, \quad \mathbb{P}\text{-a.s.}, \quad (3.4)$$

where $\hat{B} = (\hat{B}_i)_{i=1}^I$.

An admissible control for the maximizer is a product measure $\hat{\mathbb{Q}} = \prod_{i=1}^I \hat{\mathbb{Q}}_i$, where each $\hat{\mathbb{Q}}_i$ is defined on $(\Omega^i, \mathcal{F}^i, \{\mathcal{F}_t^i\})$, such that

$$\frac{d\hat{\mathbb{Q}}_i(t)}{d\mathbb{P}_i(t)} = \exp \left\{ \int_0^t \hat{\psi}_i(s) d\hat{B}_i(s) - \frac{1}{2} \int_0^t \hat{\psi}_i^2(s) ds \right\}, \quad t \in \mathbb{R}_+, \quad (3.5)$$

for an $\{\mathcal{F}_t\}$ -progressively measurable process $\hat{\psi} = (\hat{\psi}_1, \dots, \hat{\psi}_I)$ satisfying

$$\mathbb{E}^\mathbb{P} \left[\int_0^\infty e^{-\varrho s} \hat{\psi}_i^2(s) ds \right] < \infty \quad \text{and} \quad \mathbb{E}^\mathbb{P} \left[e^{\frac{1}{2} \int_0^t \hat{\psi}_i^2(s) ds} \right] < \infty \quad t \in \mathbb{R}_+, \quad i \in \mathcal{I}. \quad (3.6)$$

As in the QCP, we consider a probability space that is constructed from small probability spaces, where each one supports the processes associated with one of the classes. The Brownian motion approximates the difference $(\hat{A}_i^n - \hat{S}_i^n(T_i^n))_{i=1}^I$ up to the covariance matrix, and (3.3) is equivalent to (2.2). Condition (3.2) follows since the rejection process (in the QCP) and also $\theta^n \cdot \hat{Y}^n$ are nondecreasing, where $\theta^n := (1/\mu_1^n, \dots, 1/\mu_I^n)$. Occasionally, we refer to \hat{R} as the rejection process in the MSDG. The buffer constraint is imposed in (3.4). Pay attention that we consolidate the processes \hat{A}_i^n and $\hat{S}_i^n(T_i^n)$ into one Brownian motion. Hence, we consider only I changes of measures instead of $2I$. For this, in the cost function we consider new ambiguity parameters, ϵ_i , $i \in \mathcal{I}$, as explained in Section 3.1.1. From (2.5)–(2.13), one can see that the processes $\psi_{i,1}^n$ and $\psi_{i,2}^n$ that are involved in the Radon–Nikodym derivatives create a new drift under the new measure. This is emerging from the process $\hat{\psi}_i$ in (3.5).

Remark 3.1 (i) Notice that the structure of the information in the game is consistent with the one in the QCP. That is, the minimizer chooses a strategy and the maximizer responds to this strategy by choosing a worst case scenario. See Remark 2.1.

(ii) Given any $\{\mathcal{F}_t\}$ -progressively measurable process $\hat{\psi}$ that satisfies the conditions in (3.6), the r.h.s. of (3.5) is a martingale, and therefore, there exists a probability measure $\hat{\mathbb{Q}}$ such that $\hat{\mathbb{Q}}|_{\mathcal{F}_t}$ satisfies (3.5) for all $t \in \mathbb{R}_+$.

(iii) Equation (3.3) can alternatively be written as

$$\hat{X}(t) = \hat{x}_0 + \hat{m}t + \int_0^t \hat{\sigma} \hat{\psi}(s) ds + \hat{\sigma} \hat{B}^{\hat{\mathbb{Q}}}(t) + \hat{Y}(t) - \hat{R}(t), \quad t \in \mathbb{R}_+, \quad (3.7)$$

where $\hat{B}^{\hat{\mathbb{Q}}}(t) := \hat{B}(t) - \int_0^t \hat{\psi}(s) ds$, $t \in \mathbb{R}_+$, is an $\{\mathcal{F}_t\}$ - I -dimensional standard Brownian motion under $\hat{\mathbb{Q}}$.

Denote by $\hat{\mathcal{A}}(\hat{x}_0)$ the set of all admissible controls for the minimizer, given the initial condition \hat{x}_0 . We often abuse notation and denote $(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)$, keeping in mind that the control includes a filtered probability space. The set of all admissible controls for the maximizer is denoted by $\hat{\mathcal{Q}}(\hat{x}_0)$.

3.1.1 The cost function (MSDG)

Recall that from the estimations in (2.13), under $\hat{\mathbb{Q}}^n$, $\hat{A}_i^n - \hat{S}_i^n(T_i^n)$ is approximately a diffusion process with drift $\lambda_i^{1/2} \hat{\psi}_i^n := \lambda_i^{1/2} \hat{\psi}_{i,1}^n - \rho_i \mu_i^{1/2} \hat{\psi}_{i,2}^n$ and a diffusion coefficient $\sigma_{ii} = (2\lambda_i)^{1/2}$. The

term ρ_i is due to the convergence $T_i^n \rightarrow \rho_i$. Also, by the same arguments that lead to (3.13) below and the Taylor expansion of $\log(1+x)$, one has

$$\begin{aligned} & \frac{1}{\kappa_{i,1}} L^\varrho(\hat{\mathbb{Q}}_{i,1}^n \| \mathbb{P}_{i,1}^n) + \frac{1}{\kappa_{i,2}} L^\varrho(\hat{\mathbb{Q}}_{i,2}^n \| \mathbb{P}_{i,2}^n) \\ & \approx \mathbb{E}^{\hat{\mathbb{Q}}_{i,1}^n \times \hat{\mathbb{Q}}_{i,2}^n} \left[\int_0^\infty e^{-\varrho t} \left\{ \frac{1}{2\kappa_{i,1}} (\hat{\psi}_{i,1}^n(t))^2 + \frac{1}{2\kappa_{i,2}} \rho_i (\hat{\psi}_{i,2}^n(t))^2 \right\} dt \right]. \end{aligned} \quad (3.8)$$

Now, since the maximizer is free to choose $\hat{\psi}_{i,1}^n$ and $\hat{\psi}_{i,2}^n$, he faces the two steps optimization problem. First, to choose $\hat{\psi}_i^n(t)$ and then to solve

$$\min_{(\hat{\psi}_{i,1}^n(t), \hat{\psi}_{i,2}^n(t))} \left\{ \frac{1}{2\kappa_{i,1}} (\hat{\psi}_{i,1}^n(t))^2 + \frac{1}{2\kappa_{i,2}} \rho_i (\hat{\psi}_{i,2}^n(t))^2 : \lambda_i^{1/2} \hat{\psi}_{i,1}^n(t) - \rho_i \mu_i^{1/2} \hat{\psi}_{i,2}^n(t) = \lambda_i^{1/2} \hat{\psi}_i^n(t) \right\}.$$

The minimal value of the above equals $\frac{1}{2\epsilon_i} (\hat{\psi}_i^n(t))^2$, where

$$\epsilon_i := \frac{1}{2} (\kappa_{i,1} + \kappa_{i,2}). \quad (3.9)$$

Therefore, (3.8) can be approximated by $\frac{1}{2\epsilon_i} (\hat{\psi}_i^n(t))^2$, which gets along with (3.12) below.

Set $\epsilon = (\epsilon_i)_{i=1}^I$. The cost associated with the initial condition \hat{x}_0 and the strategies (\hat{Y}, \hat{R}) and $\hat{\mathbb{Q}}$ is given by

$$\hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \epsilon) := \mathbb{E}^{\hat{\mathbb{Q}}} \left[\int_0^\infty e^{-\varrho t} \left(\hat{h} \cdot \hat{X}(t) dt + \hat{r} \cdot d\hat{R}(t) \right) \right] - \sum_{i=1}^I \frac{1}{\epsilon_i} L^\varrho(\hat{\mathbb{Q}}_i \| \mathbb{P}_i), \quad (3.10)$$

where

$$L^\varrho(\hat{\mathbb{Q}}_i \| \mathbb{P}_i) := \mathbb{E}^{\hat{\mathbb{Q}}_i} \left[\int_0^\infty \varrho e^{-\varrho t} \log \frac{d\hat{\mathbb{Q}}_i(t)}{d\mathbb{P}_i(t)} dt \right] \quad (3.11)$$

is a discounted variant of the Kullback–Leibler divergence. The cost function can alternatively be expressed by

$$\hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \epsilon) = \mathbb{E}^{\hat{\mathbb{Q}}} \left[\int_0^\infty e^{-\varrho t} \left(\hat{h} \cdot \hat{X}(t) dt + \hat{r} \cdot d\hat{R}(t) - \sum_{i=1}^I \frac{1}{2\epsilon_i} \hat{\psi}_i^2(t) dt \right) \right], \quad (3.12)$$

with $\hat{\psi}$ satisfying (3.5)–(3.6) above. Indeed,

$$\begin{aligned} L^\varrho(\hat{\mathbb{Q}}_i \| \mathbb{P}_i) &= \mathbb{E}^{\hat{\mathbb{Q}}_i} \left[\int_0^\infty \varrho e^{-\varrho t} \left(-\frac{1}{2} \int_0^t \hat{\psi}_i^2(s) ds + \int_0^t \hat{\psi}_i(s) d\hat{B}_i(s) \right) dt \right] \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_i} \left[\int_0^\infty \varrho e^{-\varrho t} \left(-\frac{1}{2} \int_0^t |\hat{\psi}_i(s)|^2 ds + \int_0^t \hat{\psi}_i(s) \cdot (d\hat{B}_i^{\hat{\mathbb{Q}}}(s) + \hat{\psi}_i(s) ds) \right) dt \right] \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_i} \left[\int_0^\infty \varrho e^{-\varrho t} \left(\frac{1}{2} \int_0^t \hat{\psi}_i^2(s) ds \right) dt \right] = \mathbb{E}^{\hat{\mathbb{Q}}_i} \left[\frac{1}{2} \int_0^\infty e^{-\varrho t} \hat{\psi}_i^2(t) dt \right] < \infty. \end{aligned} \quad (3.13)$$

While the form of the cost function given in (3.10) captures better the ambiguity aversion, the form of the cost given in (3.12) is more useful from a technical point of view. Moreover, the dynamics in (3.7) together with the cost function given in (3.12) are similar in their structure to their correspondences in Equation (11) and the display below (13) together with (2) in [2].

The DM is faced the following robust optimization problem

$$\hat{V}(\hat{x}_0; \epsilon) = \inf_{(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)} \sup_{\hat{\mathbb{Q}} \in \hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \epsilon)$$

3.2 The reduced stochastic differential game (RSDG)

We now present a one-dimensional stochastic differential game to be referred to as the *reduced stochastic differential game* (RSDG). This game is obtained by projecting the processes from (3.3) in the $\theta = (\mu_i^{-1})_{i=1}^I$ direction, which is given in (3.2). For this we need the following notation,

$$x_0 := \theta \cdot \hat{x}_0, \quad m := \theta \cdot \hat{m}, \quad \sigma := \|\theta \hat{\sigma}\|, \quad (3.14)$$

$$\varepsilon := \frac{1}{\sigma^2} \sum_{i=1}^I (\theta \hat{\sigma})_i^2 \epsilon_i, \quad (3.15)$$

and

$$b := \max\{\theta \cdot \hat{\xi} : \hat{\xi} \in \mathcal{X}\} = \sum_{i=1}^I \mu_i^{-1} b_i.$$

Definition 3.2 (admissible controls, RSDG) *An admissible control for the minimizer for any initial state $x_0 \in [0, b]$ is a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ that supports a one-dimensional standard Brownian motion B and a process (Y, R) taking values in \mathbb{R}_+^2 with RCLL sample paths, both adapted to the filtration $\{\mathcal{F}_t\}$ and satisfy the following properties:*

- (i) *for every $0 \leq s < t$, $B(t) - B(s)$ is independent of \mathcal{F}_s under \mathbb{P} ;*
- (ii) *Y and R are nonnegative and nondecreasing;*
- (iii)

$$X(t) = x_0 + mt + \sigma B(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+, \quad (3.16)$$

such that

$$X(t) \in [0, b], \quad t \in \mathbb{R}_+, \quad \mathbb{P}\text{-a.s.}$$

An admissible control for the maximizer is a measure \mathbb{Q} defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\})$ such that

$$\frac{d\mathbb{Q}(t)}{d\mathbb{P}(t)} = \exp \left\{ \int_0^t \psi(s) dB(s) - \frac{1}{2} \int_0^t \psi^2(s) ds \right\}, \quad t \in \mathbb{R}_+, \quad (3.17)$$

for an $\{\mathcal{F}_t\}$ -progressively measurable process ψ satisfying

$$\mathbb{E}^{\mathbb{P}} \left[\int_0^\infty e^{-\varrho s} \psi^2(s) ds \right] < \infty \quad \text{and} \quad \mathbb{E}^{\mathbb{P}} \left[e^{\frac{1}{2} \int_0^t \psi^2(s) ds} \right] < \infty \quad \text{for every } t \in \mathbb{R}_+. \quad (3.18)$$

The statements given in Remark 3.1 also hold for the RSDG as well. For completeness of the presentation and for later references we provide an alternative form of the dynamics given in (3.16),

$$X(t) = x_0 + mt + \int_0^t \sigma \psi(s) ds + \sigma B^\mathbb{Q}(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+, \quad (3.19)$$

where $B^\mathbb{Q}(t) := B(t) - \int_0^t \psi(s) ds$, $t \in \mathbb{R}_+$, is an $\{\mathcal{F}_t\}$ -one-dimensional standard Brownian motion under \mathbb{Q} .

Denote by $\mathcal{A}(x_0)$ the set of all admissible controls for the minimizer, given the initial condition x_0 . As before, we often abuse notation and denote $(Y, R) \in \mathcal{A}(x_0)$, keeping in mind that the control includes a filtered probability space. The set of all admissible controls for the maximizer is denoted by $\mathcal{Q}(x_0)$.

3.2.1 The cost function (RSDG)

The cost associated with the initial condition x and the controls (Y, R) and \mathbb{Q} is given by

$$J(x_0, Y, R, \mathbb{Q}; \varepsilon) := \mathbb{E}^\mathbb{Q} \left[\int_0^\infty e^{-\varrho t} (h(X(t)) dt + r dR(t)) \right] - \frac{1}{\varepsilon} L^\varrho(\mathbb{Q} \| \mathbb{P}),$$

where

$$h(x) := \min \{ \hat{h} \cdot \hat{\xi} : \hat{\xi} \in \mathcal{X}, \theta \cdot \hat{\xi} = x \}, \quad (3.20)$$

$$r := \min \{ \hat{r} \cdot q : q \in \mathbb{R}_+^2, \theta \cdot q = 1 \}, \quad (3.21)$$

and $L^\varrho(\mathbb{Q} \| \mathbb{P})$ is given by (3.11) with (\mathbb{Q}, \mathbb{P}) replacing $(\hat{\mathbb{Q}}_i, \mathbb{P}_i)$. By the convexity of \mathcal{X} it follows that h is convex. In fact, h is piecewise linear and Lipschitz continuous. Moreover, $h(x) \geq 0$ for $x \geq 0$ and equality holds if and only if $x = 0$. Therefore, h is strictly increasing. In [6, page 568] it is shown that there is $i^* \in \mathcal{I}$ such that,

$$r = r_{i^*} \mu_{i^*} := \min \{ r_i \mu_i : i \in \mathcal{I} \}. \quad (3.22)$$

The index i^* stands for the class with the lowest rejection cost. In fact, as we discuss in Section 3.3 and prove in Theorem 5.1, under optimality of both players in the MSDG, rejections are performed only from this class.

By the same arguments that lead to (3.12), the cost function of the RSDG can alternatively be expressed by the technically more convenient form,

$$J(x_0, Y, R, \mathbb{Q}; \varepsilon) = \mathbb{E}^\mathbb{Q} \left[\int_0^\infty e^{-\varrho t} \left(h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi^2(t) dt \right) \right], \quad (3.23)$$

with ψ satisfying (3.18) above. The value function is given by

$$V(x_0; \varepsilon) = \inf_{(Y, R) \in \mathcal{A}(x_0)} \sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} J(x_0, Y, R, \mathbb{Q}; \varepsilon). \quad (3.24)$$

Remark 3.2 *In case that there is no ambiguity, we define the cost and the value functions by*

$$J_{NA}(x_0, Y, R) := \mathbb{E} \left[\int_0^\infty e^{-\rho t} (h(X(t))dt + \rho dR(t)) \right], \quad (3.25)$$

$$V(x_0; 0) := \inf_{(Y, R) \in \mathcal{A}(x_0)} J_{NA}(x_0, Y, R).$$

This problem was studied by Harrison and Taksar [18] and later on was used by Atar and Shifrin [6]. In Theorem 6.1 we show that this problem is obtained when the ambiguity vanishes, that is, $\lim_{\varepsilon \rightarrow 0} V(x_0; \varepsilon) = V(x_0; 0)$.

3.3 The relationship between the games

We now show that the last two games share the same value and moreover, that given any admissible control for the minimizer in either one of the games, one can construct an admissible control in the other game that performs at least as well. To this end, we define a function γ , taken from [6, Equations (48)–(49)], that sends any workload value to the cheapest state of the MSDG from the holding cost perspective, for which the workload is the given one. Using this function and an optimal strategy for the minimizer in the RSDG, we construct an optimal strategy for the minimizer in the MSDG, see Theorem 5.1. In order to define the function it is convenient to assume without loss of generality that

$$h_1 \mu_1 \geq h_2 \mu_2 \geq \dots \geq h_I \mu_I.$$

Recall that $b = \theta \cdot (b_1, \dots, b_I)$. Given $x_0 \in [0, b]$, let (j, v) be the unique pair that is determined by

$$x_0 = \sum_{i=j+1}^I \theta_i b_i + \theta_j v, \quad j \in \mathcal{I}, \quad v \in [0, b_j],$$

and for $x_0 = b$, take $(j, v) = (1, b_1)$. Let $\gamma : [0, b] \rightarrow \mathcal{X}$ be the function given by

$$\gamma(x) = \sum_{i=j+1}^I b_i e_i + v e_j, \quad (3.26)$$

where $\{e_1, \dots, e_I\}$ is the standard basis of \mathbb{R}^I . The curve $\gamma(x)$, $x \in [0, b]$ is continuous and located on the edges of \mathcal{X} , see Figure 1. The idea is as follows, recall that the components of $\hat{Y} = (\hat{Y}_i)_{i=1}^I$ can be positive or negative, as long as $\theta \cdot \hat{Y}$ is nonnegative and nondecreasing. Now, as the workload changes in the interval $[0, b]$, the DM can use only the process \hat{Y} , without the need of the rejection process \hat{R} , so that \hat{X} moves along the curve of γ . As will be shown in Theorem 5.1, under optimality, the rejection process is used only to reduce the workload, and only from the class which has the cheapest rejections cost, denoted by i^* . We discuss more about the minimizer's optimal strategy in the MSDG in Remark 3.3 and in the paragraph that comes before theorem 5.1.

Since our starting point in this paper is the QCP, we state the proposition below for an arbitrary initial point $\hat{x}_0 \in \mathcal{X}$ and show that $\hat{V}(\hat{x}_0; \varepsilon) = V(x_0; \varepsilon)$, where recall that $x_0 = \theta \cdot \hat{x}_0$.

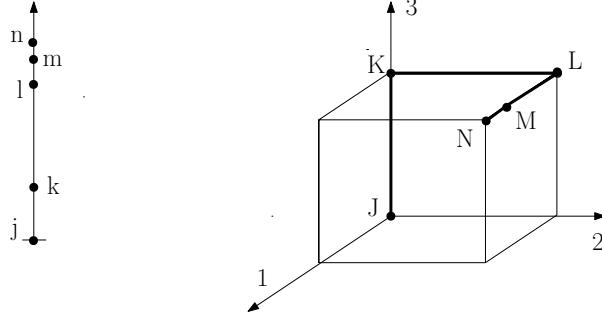


Figure 1: The graphs refer to the case $I = 3$, $\hat{h} = (1, 5/2, 3/2)$, $\mu = (3, 1, 3/2)$, and $(b_1, b_2, b_3) = (4, 7, 6)$. The graph to the left stands for the workload levels. The curve of the function γ is in bold in the graph to the right. The workload levels with the lower case letters are $j = 0, k = b_3/\mu_3 = 4, l = b_3/\mu_3 + b_2/\mu_2 = 11, m = b_3/\mu_3 + b_2/\mu_2 + 1$, and $n = b_3/\mu_3 + b_2/\mu_2 + b_1/\mu_1 = 37/3$. They respectively correspond to the upper case letters: $J = (0, 0, 0), K = (0, 0, b_3/\mu_3) = (0, 0, 4), L = (0, b_2/\mu_2, b_3/\mu_3) = (0, 7, 4), M = (1, b_2/\mu_2, b_3/\mu_3) = (1, 7, 4)$, and $N = (b_1/\mu_1, b_2/\mu_2, b_3/\mu_3) = (4/3, 7, 4)$.

Proposition 3.1 Fix $\epsilon = (\epsilon_i)_{i=1}^I$ and $\hat{x}_0 \in \mathcal{X}$. Let ε be given by (3.15).

(i) Given an admissible control in the MSDG, $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, \hat{B}, \hat{Y}, \hat{R})$ and an admissible measure \mathbb{Q}_* associated with ψ_* that satisfies (3.17)–(3.18), set $B = \frac{1}{\sigma}\theta\hat{\sigma} \cdot \hat{B}$, $(X, Y, R) = (\theta \cdot \hat{X}, \theta \cdot \hat{Y}, \theta \cdot \hat{R})$, and $\hat{\psi}_* = (\hat{\psi}_{*,1}, \dots, \hat{\psi}_{*,I})$, by

$$\hat{\psi}_{*,i}(t) := \frac{\sigma\psi_*(t)(\theta\hat{\sigma})_i\epsilon_i}{\sum_{j=1}^I(\theta\hat{\sigma})_j^2\epsilon_j}, \quad t \in \mathbb{R}_+, i \in \mathcal{I}. \quad (3.27)$$

Let $\hat{\mathbb{Q}}_*$ be the associated measure defined through (3.5). Then $(Y, R) \in \mathcal{A}(x_0)$, $\hat{\mathbb{Q}}_* \in \hat{\mathcal{Q}}(\hat{x}_0)$, and

$$J(x_0, Y, R, \mathbb{Q}_*; \varepsilon) \leq \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_*; \varepsilon). \quad (3.28)$$

As a consequence,

$$\sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} J(x_0, Y, R, \mathbb{Q}; \varepsilon) \leq \sup_{\hat{\mathbb{Q}} \in \hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \varepsilon).$$

(ii) Conversely, consider an admissible control in the RSDG, $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, B, Y, R)$, which is assumed to support an I -dimensional standard Brownian motion \hat{B} . Consider also an admissible $\hat{\mathbb{Q}}_\#$ associated with $\hat{\psi}_\#$ that satisfies (3.5)–(3.6). Assume that \hat{B} is $\{\mathcal{F}_t\}$ -adapted and satisfies $\frac{1}{\sigma}\theta\hat{\sigma} \cdot \hat{B} = B$ and (3.1). Define $(\hat{X}, \hat{Y}, \hat{R})$ by

$$\hat{X}(t) := \gamma(X(t)), \quad \hat{R}(t) := R(t)\mu_{i^*}e_{i^*}, \quad (3.29)$$

and

$$\hat{Y}(t) := \hat{X}(t) - \hat{x}_0 - \hat{m}t - \hat{\sigma}\hat{B}(t) + \hat{R}(t). \quad (3.30)$$

Also, let

$$\psi_\#(t) := \frac{1}{\sigma}\theta\hat{\sigma} \cdot \hat{\psi}_\#(t), \quad t \in \mathbb{R}_+, \quad (3.31)$$

and let $\mathbb{Q}_\#$ be the associated measure defined through (3.17). Then $(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)$, $\mathbb{Q}_\# \in \mathcal{Q}(x_0)$, and

$$\hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_\#; \epsilon) \leq J(x_0, Y, R, \mathbb{Q}_\#; \epsilon).$$

As a consequence,

$$\sup_{\hat{\mathbb{Q}} \in \hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \epsilon) \leq \sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} J(x_0, Y, R, \mathbb{Q}; \epsilon). \quad (3.32)$$

Corollary 3.1 Fix $\epsilon = (\epsilon_i)_{i=1}^I$ and $\hat{x}_0 \in \mathcal{X}$. Let ε be given by (3.15). Then, $\hat{V}(\hat{x}_0; \epsilon) = V(x_0; \varepsilon)$.

The Corollary follows by Proposition 3.1 once it is proved that the probability space in part (ii) of the proposition supports an I -dimensional standard Brownian motion \hat{B} . The proof of this argument is merely technical and is fully given in [6, Proposition 2.1.(iii)]. Therefore, it is omitted.

Remark 3.3 In Theorem 4.1 we show that the minimizer in the RSDG has a simple optimal strategy, which uses minimal idleness and minimal amount of rejections in order to keep the workload in some subinterval $[0, \beta] \subseteq [0, b]$. In Theorem 5.1 we use the function γ given in (3.26) and the relations given in (3.29)–(3.30) to construct an optimal strategy for the minimizer in the MSDG. An optimal strategy for the maximizer emerges from (3.27) and (3.31).

Proof of Proposition 3.1: Some of the arguments in the proof are given in [6, Proposition 2.1]. However, the arguments that involve changes of the probability measure are new. For completeness of the proof we provide all the details.

(i) The proof that $(Y, R) \in \mathcal{A}(x_0)$ is straightforward and therefore omitted. Notice that $\|\hat{\psi}_*(\cdot)\| \leq |\psi_*(\cdot)|$ and therefore, $\hat{\mathbb{Q}}_* \in \hat{\mathcal{Q}}(\hat{x}_0)$ follows since $\mathbb{Q}_* \in \mathcal{Q}(x_0)$.

We now show that the distribution of X under \mathbb{Q}_* is the same as under $\hat{\mathbb{Q}}_*$. Replacing $\hat{\mathbb{Q}}$ by $\hat{\mathbb{Q}}_*$ in equation (3.7) and multiplying its both sides by θ yield

$$\begin{aligned} X(t) &= x_0 + mt + \int_0^t \theta \hat{\sigma} \hat{\psi}_*(s) ds + \theta \hat{\sigma} \hat{B}^{\hat{\mathbb{Q}}_*}(t) + Y(t) - R(t) \\ &= x_0 + mt + \int_0^t \sigma \psi_*(s) ds + \sigma B^{\mathbb{Q}_*}(t) + Y(t) - R(t). \end{aligned} \quad (3.33)$$

The equality between the integrals follows by the definitions of $\hat{\psi}_*$ and σ . The equality between the Brownian motion terms follows since

$$\theta \hat{\sigma} \hat{B}^{\hat{\mathbb{Q}}_*}(t) = \theta \hat{\sigma} \hat{B}(t) - \int_0^t \theta \hat{\sigma} \hat{\psi}_*(s) ds = \sigma B(t) - \sigma \int_0^t \psi_*(s) ds = \sigma B^{\mathbb{Q}_*},$$

which in turn follows by using the definitions of $\hat{B}^{\hat{\mathbb{Q}}_*}$ and $B^{\mathbb{Q}_*}$ ((3.7) and (3.19)) together with the definition of B from the proposition and once again the definition of $\hat{\psi}_*$. Recall that $\hat{B}^{\hat{\mathbb{Q}}_*}$

is an I -dimensional standard Brownian motion under $\hat{\mathbb{Q}}_*$ and that $\theta\hat{\sigma}$ is a deterministic vector with norm σ . Then from the above, we get that under $\hat{\mathbb{Q}}_*$, the process $B^{\mathbb{Q}_*}$ is a one-dimensional standard Brownian motion. Since $B^{\mathbb{Q}_*}$ is a one-dimensional standard Brownian motion also under \mathbb{Q}_* , we get from (3.33) that X admits the same distribution under $\hat{\mathbb{Q}}_*$ and under \mathbb{Q}_* .

Next, by the definitions of h and r (see (3.20) and (3.21)) it follows that

$$h(X(t)) \leq \hat{h} \cdot \hat{X}(t), \quad t \in \mathbb{R}_+, \quad (3.34)$$

$$\int_0^\infty e^{-\varrho t} r dR(t) \leq \int_0^\infty e^{-\varrho t} \hat{r} \cdot d\hat{R}(t). \quad (3.35)$$

Moreover, using the fact that the distribution of X under $\hat{\mathbb{Q}}_*$ is the same as under \mathbb{Q}_* and the equality

$$\sum_{i=1}^I \frac{1}{2\epsilon_i} \hat{\psi}_{*,i}^2(t) = \frac{1}{2\epsilon} \psi_*^2, \quad t \in \mathbb{R}_+,$$

we get,

$$\begin{aligned} J(x_0, Y, R, \mathbb{Q}_*; \epsilon) &= \mathbb{E}^{\mathbb{Q}_*} \left[\int_0^\infty e^{-\varrho t} \left(h(X(t)) dt + r dR(t) - \frac{1}{2\epsilon} \psi_*^2(t) \right) dt \right] \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_*} \left[\int_0^\infty e^{-\varrho t} \left(h(X(t)) dt + r dR(t) - \frac{1}{2\epsilon} \psi_*^2(t) \right) dt \right] \\ &\leq \mathbb{E}^{\hat{\mathbb{Q}}_*} \left[\int_0^\infty e^{-\varrho t} \left(\hat{h} \cdot \hat{X}(t) dt + \hat{r} \cdot d\hat{R}(t) - \sum_{i=1}^I \frac{1}{2\epsilon_i} \hat{\psi}_{*,i}^2(t) \right) dt \right] \\ &= \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_*; \epsilon). \end{aligned}$$

Since $\mathbb{Q}_* \in \mathcal{Q}(x_0)$ is arbitrary, it follows that

$$\sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} J(x_0, Y, R, \mathbb{Q}; \epsilon) \leq \sup_{\hat{\mathbb{Q}} \in \hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \epsilon).$$

(ii) We start by showing that $(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)$. The processes \hat{X} , \hat{R} , and \hat{Y} are $\{\mathcal{F}_t\}$ -adapted since X , R , and \hat{B} are. Now,

$$\theta \cdot \hat{Y}(t) = \theta \cdot \hat{X}(t) - \theta \cdot x_0 - \theta \cdot \hat{m}t - \theta \cdot \hat{\sigma} \hat{B}(t) + \theta \cdot \hat{R}(t) = Y(t),$$

and since Y is admissible it follows that $\theta \cdot \hat{Y} = Y$ is nonnegative and nondecreasing. The processes $\{\hat{R}_i\}_{i=1}^I$ are clearly nonnegative and nondecreasing since so is R . Thus, property (3.2) holds. Finally, properties (3.3) and (3.4) follow by the definition of γ and the construction of $(\hat{X}, \hat{Y}, \hat{R})$.

Notice that, $\|\hat{\psi}_\#(\cdot)\| = |\psi_\#(\cdot)|$ and therefore, $\mathbb{Q}_\# \in \mathcal{Q}(x_0)$ follows since $\hat{\mathbb{Q}}_\# \in \hat{\mathcal{Q}}(\hat{x}_0)$. From Cauchy-Schwartz inequality,

$$\sum_{i=1}^I \frac{1}{2\epsilon_i} \hat{\psi}_{\#,i}^2(t) \geq \frac{1}{2\epsilon} \psi_\#^2(t).$$

By the definitions of h , r , and γ it follows that

$$\begin{aligned}\hat{h} \cdot \hat{X}(t) &= h(X(t)), \quad t \in \mathbb{R}_+, \\ \int_0^\infty e^{-\varrho t} \hat{r} \cdot d\hat{R}(t) &= \int_0^\infty e^{-\varrho t} r d(R(t)).\end{aligned}$$

By the relationship between the I -dimensional processes \hat{B} and $\hat{\psi}_\#$ and the one-dimensional processes B and $\psi_\#$, one can show as was done in the previous part that the distribution of X under $\hat{\mathbb{Q}}_\#$ is the same as under $\mathbb{Q}_\#$. Therefore,

$$\begin{aligned}J(x_0, Y, R, \mathbb{Q}_\#; \varepsilon) &= \mathbb{E}^{\mathbb{Q}_\#} \left[\int_0^\infty e^{-\varrho t} \left(h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi_\#^2(t) \right) dt \right] \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_\#} \left[\int_0^\infty e^{-\varrho t} \left(h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi_\#^2(t) \right) dt \right] \\ &\geq \mathbb{E}^{\hat{\mathbb{Q}}_\#} \left[\int_0^\infty e^{-\varrho t} \left(\hat{h} \cdot \hat{X}(t) dt + \hat{r} \cdot d\hat{R}(t) - \sum_{i=1}^I \frac{1}{2\varepsilon_i} \hat{\psi}_{\#,i}^2(t) \right) dt \right] \\ &= \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_\#; \varepsilon).\end{aligned} \tag{3.36}$$

Since $\hat{\mathbb{Q}}_\# \in \hat{\mathcal{Q}}(\hat{x}_0)$ is arbitrary, it follows that

$$\sup_{\hat{\mathbb{Q}} \in \hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \varepsilon) \leq \sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} J(x_0, Y, R, \mathbb{Q}; \varepsilon).$$

□

4 Solution of the WBCP

In this section we provide a solution to the RSDG. In Section 4.1 we present the notion of a reflection strategy. Then in Section 4.2 we provide the HJB equation associated with the RSDG. We prove that the value function is the unique smooth solution of the HJB equation and show that the minimizer has an optimal reflecting strategy.

4.1 Reflecting strategies

The optimal strategy of the minimizer is shown to be one that enforces the workload to stay in a specific interval of the form $[0, \beta]$ with minimal effort. To rigorously define such a strategy we make use of the *Skorokhod map on an interval*. Fix $\alpha < \beta$. For any $\eta \in \mathcal{D}(\mathbb{R}_+, \mathbb{R})$ there exists a unique triplet of functions $(\chi, \zeta_1, \zeta_2) \in \mathcal{D}(\mathbb{R}_+, \mathbb{R}^3)$ that satisfies the following properties:

- (i) for every $t \in \mathbb{R}_+$, $\chi(t) = \eta(t) + \zeta_1(t) - \zeta_2(t)$;
- (ii) ζ_1 and ζ_2 are nondecreasing, $\zeta_1(0-) = \zeta_2(0-) = 0$, and

$$\int_0^\infty 1_{(\alpha, \beta]}(\chi(t)) d\zeta_1(t) = \int_0^\infty 1_{[\alpha, \beta)}(\chi(t)) d\zeta_2(t) = 0.$$

We denote by $\Gamma_{[\alpha, \beta]}(\eta) = (\Gamma_{[\alpha, \beta]}^1, \Gamma_{[\alpha, \beta]}^2, \Gamma_{[\alpha, \beta]}^3)(\eta) = (\chi, \zeta_1, \zeta_2)$. See [22] for existence and uniqueness of solutions, and continuity and further properties of the map. In particular, we have the following.

Lemma 4.1 *There exists a constant $c_S > 0$ such that for every $t > 0$, $\alpha < \beta$ and $\omega, \tilde{\omega} \in \mathcal{D}(\mathbb{R}_+, \mathbb{R})$,*

$$\sup_{s \in [0, t]} \|\Gamma_{[\alpha, \beta]}(\omega)(t) - \Gamma_{[\alpha, \beta]}(\tilde{\omega})(t)\| \leq c_S \sup_{s \in [0, t]} |\omega(s) - \tilde{\omega}(s)|.$$

Definition 4.1 *Fix $x_0, \beta \in [0, b]$. The strategy (Y, R) is called a β -reflecting strategy if for every $\eta \in \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ one has $(X, Y, R)(\eta) = \Gamma_{[0, \beta]}(\eta)$.*

One can easily verify that any β -reflecting strategy is admissible.

4.2 The HJB equation and the value function

In case that there is no ambiguity (see Remark 3.2), the problem was analyzed by Harrison and Taksar [18]. It is shown there (see Proposition 5.11) that if $k_{\alpha\beta}$ is a $\mathcal{C}^1([0, b], \mathbb{R})$ is twice continuously differentiable on $[\alpha, \beta]$ and satisfies

$$l + k'_{\alpha\beta}(x) = 0, \quad 0 \leq x \leq \alpha, \quad (4.1)$$

$$\frac{1}{2}\sigma^2 k''_{\alpha\beta}(x) + mk'_{\alpha\beta}(x) - \varrho k_{\alpha\beta}(x) + h(x) = 0, \quad \alpha \leq x \leq \beta, \quad (4.2)$$

$$r - k'_{\alpha\beta}(x) = 0, \quad \beta \leq x \leq b, \quad (4.3)$$

then

$$k_{\alpha\beta}(x) = \mathbb{E}^\mathbb{P} \left[\int_0^\infty e^{-\varrho t} [h(X(t))dt + r dR(t) + l dY(t)] \right], \quad x \in [0, b],$$

with $(X, Y, R)(t) = \Gamma_{[\alpha, \beta]}(x_0 + m \cdot + \sigma B(\cdot))(t)$, $t \in \mathbb{R}_+$. The rational behind this argument is as follows. Consider $x \in (\beta, b)$, then in order to keep the process X between α and β , there is an instantaneous reflection from above that contributes the cost $r(x - \beta)$. This explains (4.3). Similar arguments yield (4.1). When no control is taking action, standard arguments imply (4.2). The HJB in this case takes the form

$$\begin{cases} \left[\frac{1}{2}\sigma^2 f''(x) + mf'(x) - \varrho f(x) + h(x) \right] \wedge f'(x) \wedge [r - f'(x)] = 0, & x \in (0, b), \\ f'(0) = 0, & f'(b) = r. \end{cases}$$

However, the uniqueness of a solution of the HJB is not argued in [18] (see the paragraph before Section 7) but rather in [6, Proposition 2.2], using viscosity solutions.

Remark 4.1 *Notice that in [18] there is a cost associated with the reflections from both sides, unlike our case that does not consider a cost component for reflections from below, that is $l = 0$. It simply follows since in our QCP there is no penalty/reward for idleness. Henceforth, under optimality, $\alpha = 0$ in (4.1)–(4.2).*

Motivated by these results and the structure of the cost function given in (3.23) together with the inf-sup structure of the value function given in (3.24), we consider the following HJB,

$$\begin{cases} \left[\sup_{p \in \mathbb{R}} \left\{ \frac{1}{2}\sigma^2 f''(x) + (m + \sigma p)f'(x) - \varrho f(x) + h(x) - \frac{1}{2\varepsilon} p^2 \right\} \right] \wedge f'(x) \wedge [r - f'(x)] = 0, & x \in (0, b), \\ f'(0) = 0, & f'(b) = r. \end{cases}$$

Or equivalently, by substituting the optimal solution of the $\sup_{p \in \mathbb{R}}$ above, $p^* = \varepsilon \sigma f'(x)$,

$$\begin{cases} [f''(x) + H(x, f(x), f'(x))] \wedge f'(x) \wedge [r - f'(x)] = 0, & x \in (0, b), \\ f'(0) = 0, & f'(b) = r, \end{cases} \quad (\text{HJB}(\varepsilon))$$

where hereafter,

$$H(x, y, z) := \frac{2}{\sigma^2} \left(mz + \frac{1}{2} \sigma^2 \varepsilon z^2 - \varrho y + h(x) \right).$$

Notice that when $\varepsilon = 0$, $\text{HJB}(\varepsilon)$ coincides with the HJB given in Harrison and Taksar [18, Equation (1.2)] and by Atar and Shifrin [6, Equation (41)].

The idea is that we may think of the change of measure done by the maximizer as a change of the drift term from m to $m + \sigma p$, which costs him $\frac{1}{2\varepsilon} p^2$, where p depends on x and is chosen in order to maximize the cost. As can be seen, due to the term $\frac{1}{2} \sigma^2 \varepsilon (f'(x))^2$ in $H(x, f(x), f'(x))$, the HJB is not linear, a fact that raises some technical difficulties in proving Proposition 4.3 below.

We now state the main theorem of the paper, which is given also for $\varepsilon = 0$, see Remark 3.2.

Theorem 4.1 *Fix $\varepsilon \in [0, \infty)$. The value function $V(\cdot; \varepsilon)$ is the unique $\mathcal{C}^2([0, b], \mathbb{R})$ solution of $\text{HJB}(\varepsilon)$. Moreover, set*

$$\beta_\varepsilon = \inf \{x \in (0, b] : V'(x; \varepsilon) = r\}, \quad (4.4)$$

where $V'(x; \varepsilon)$ is the derivative of $V(\cdot; \varepsilon)$ w.r.t. x . Then the β_ε -reflecting strategy is optimal for the minimizer and $V = V(\cdot; \varepsilon)$ satisfies,

$$\begin{cases} V''(x) + H(x, V(x), V'(x)) = 0, & 0 \leq x \leq \beta_\varepsilon, \\ r - V'(x) = 0, & \beta_\varepsilon \leq x \leq b, \\ V'(0) = 0. \end{cases} \quad (4.5)$$

From the definition of $\text{HJB}(\varepsilon)$ and the theorem above we get the following corollary, which is given for reference purposes.

Corollary 4.1 *For any $\varepsilon \in (0, \infty)$, $0 \leq V'(\cdot; \varepsilon) \leq r$.*

The proof of the theorem is done in several steps and follows from the next three propositions. Before stating them we define a measure \mathbb{Q}_f associated with a processes ψ_f , which in turn is driven by a function $f \in \mathcal{C}^1([0, b], \mathbb{R})$ through (3.17). This measure serves us in the sequel, especially in Propositions 4.1 and 4.2 and Theorem 5.1. In the latter, we show that the measure $Q_V = Q_{V(\cdot; \varepsilon)}$ is the optimal strategy of the maximizer, where V is the value function. For any $f \in \mathcal{C}^1([0, b], \mathbb{R})$ and $t \in \mathbb{R}_+$ set

$$\begin{aligned} \psi_f(t) &:= \arg \max_{p \in \mathbb{R}} \left\{ \frac{1}{2} \sigma^2 f''(X(t)) + (m + \sigma p) f'(X(t)) - \varrho f(X(t)) + h(X(t)) - \frac{1}{2\varepsilon} p^2 \right\} \\ &= \varepsilon \sigma f'(X(t)). \end{aligned} \quad (4.6)$$

This is an $\{\mathcal{F}_t\}$ -progressively measurable process since Y, R , and B are, see Definition 3.2. Let \mathbb{Q}_f be the measure associated with ψ_f through (3.17). Note that ψ_f is bounded since $f \in \mathcal{C}^1([0, b], \mathbb{R})$ and therefore (3.18) holds trivially.

The proof of the propositions below are deferred to after the proof of Theorem 4.1.

Proposition 4.1 *Fix $\varepsilon \in (0, \infty)$. Assume that $HJB(\varepsilon)$ admits a $\mathcal{C}^2([0, b], \mathbb{R})$ solution f . Then*

$$f(x) \leq \inf_{(Y, R) \in \mathcal{A}(x)} J(x, Y, R, \mathbb{Q}_f; \varepsilon), \quad x \in [0, b],$$

and as a consequence $f \leq V$.

The next proposition characterizes the solution of the following ordinary differential equation,

$$\begin{cases} k''_\beta(x) + H(x, k_\beta(x), k'_\beta(x)) = 0, & 0 \leq x \leq \beta, \\ r - k'_\beta(x) = 0, & \beta \leq x \leq b, \\ k'_\beta(0) = 0. \end{cases} \quad (4.7)$$

Proposition 4.2 *Fix $\varepsilon \in (0, \infty)$. Assume that there is a function $k_\beta = k_{\beta, \varepsilon} \in \mathcal{C}^1([0, b], \mathbb{R}) \cap \mathcal{C}^2([a, b] \setminus \{\beta\}, \mathbb{R})$ that solves (4.7). Let (Y_β, R_β) be a β -reflecting strategy, that is $(X, Y_\beta, R_\beta)(t) = \Gamma_{[0, \beta]}(x + m \cdot + \sigma B(\cdot))(t)$, $t \in \mathbb{R}_+$. Then, for every $x \in [0, b]$,*

$$k_\beta(x) = \sup_{\mathbb{Q} \in \mathcal{Q}(x)} J(x, Y_\beta, R_\beta, \mathbb{Q}; \varepsilon) = J(x, Y_\beta, R_\beta, \mathbb{Q}_k; \varepsilon), \quad (4.8)$$

where $\mathbb{Q}_k = \mathbb{Q}_{k_\beta}$ is the measure associated with $\psi_k = \psi_{k_\beta}$, defined in (4.6).

We now claim that for every $\varepsilon \in (0, \infty)$, $HJB(\varepsilon)$ admits a unique smooth solution. Harrison and Taksar provided in [18, page 450] explicit functions from which a smooth solution can be constructed. The construction is provided and the smoothness is claimed to be straightforward yet tedious and therefore omitted, see the paragraph below (6.8) there. The situation is more subtle in our case since $HJB(\varepsilon)$ is nonlinear and therefore explicit solutions are out of reach. We choose a different path and use the *shooting method* to prove that a smooth solution of $HJB(\varepsilon)$ uniquely exists and that it solves the free-boundary problem (4.7). In short, the shooting method is used to solve boundary value problems by reducing them to initial value problems; see [29, Section 7.3] for further reading about the method. We take it one step forward and use it in the free-boundary setup with Neumann boundary conditions. From the technical point of view, the proof of Proposition 4.3 is the most demanding in this section.

Proposition 4.3 *For every $\varepsilon \in (0, \infty)$, $HJB(\varepsilon)$ admits a unique $\mathcal{C}^2([0, b], \mathbb{R})$ solution. Moreover, the solution wears the form k_{β_ε} for some parameter $\beta_\varepsilon \in (0, b]$ for which $(k_{\beta_\varepsilon})' < r$ on $[0, \beta_\varepsilon)$, where k_{β_ε} is taken from Proposition 4.2.*

Proof of Theorem 4.1: Recall that when $\varepsilon = 0$, $HJB(\varepsilon)$ coincides with the HJB given in [18, 6]. Thus, we focus only on positive ε 's.

From Proposition 4.3, the $HJB(\varepsilon)$ admits a unique $\mathcal{C}^2([0, b], \mathbb{R})$ solution that also solves (4.7) for some $\beta_\varepsilon \in (0, b]$. Denote it by k_{β_ε} . From Proposition 4.2 this is the cost of the β_ε -reflecting strategy. Therefore, the value function, which is the infimum over all the strategies

satisfies, $V(\cdot; \varepsilon) \leq k_{\beta_\varepsilon}(\cdot)$. Together with Proposition 4.1, $V(\cdot; \varepsilon) = k_{\beta_\varepsilon}(\cdot)$. Recalling again that k_{β_ε} is the cost of the β_ε -reflecting strategy, we obtain its optimality. The relation in (4.4) follows now from Proposition 4.3. \square

The rest of the section is devoted to the proofs of Propositions 4.1–4.3.

Proof of Proposition 4.1: Fix $f \in \mathcal{C}^2([0, b], \mathbb{R})$ and an arbitrary $(Y, R) \in \mathcal{A}(x)$ with an associated standard Brownian motion B . Set

$$X(t) = x + mt + \sigma B(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+.$$

Recalling (3.19), Itô's lemma implies that for every $t > 0$ and every $\mathbb{Q} \in \mathcal{Q}(x)$,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} [e^{-\varrho t} f(X(t))] &= f(x) + \mathbb{E}^{\mathbb{Q}} \left[\int_0^t e^{-\varrho s} \left(\frac{1}{2} \sigma^2 f''(X(s)) + (m + \sigma \psi(s)) f'(X(s)) - \varrho f(X(s)) \right) ds \right] \\ &\quad + \mathbb{E}^{\mathbb{Q}} \left[\int_0^t e^{-\varrho s} f'(X(s)) (dY^c(s) - dR^c(s)) \right] \\ &\quad + \mathbb{E}^{\mathbb{Q}} \left[\sum_{0 \leq s \leq t} e^{-\varrho s} \Delta f(X)(s) \right], \end{aligned} \quad (4.9)$$

where \mathbb{Q} and ψ are related to each other through (3.17)–(3.18). We used the fact that $B^{\mathbb{Q}}(t) := B(t) - \int_0^t \psi(s) ds$, $t \in \mathbb{R}_+$ is a \mathbb{Q} standard Brownian motion. The processes

$$Y^c(t) := Y(t) - \sum_{0 \leq s \leq t} \Delta Y(s), \quad R^c(t) := R(t) - \sum_{0 \leq s \leq t} \Delta R(s), \quad t \in \mathbb{R}_+, \quad (4.10)$$

are the continuous parts of Y and R , respectively. Consider now (4.9) with $\psi = \psi_f$ defined in (4.6) and with the measure $\mathbb{Q} = \mathbb{Q}_f$. Recalling the definition of ψ_f and that f solves HJB(ε), we get that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}_f} [e^{-\varrho t} f(X(t))] &\geq f(x) - \mathbb{E}^{\mathbb{Q}_f} \left[\int_0^t e^{-\varrho s} \left(h(X(s)) ds + r dR(s) - \frac{1}{2\varepsilon} \psi_f^2(s) ds \right) \right] \\ &\quad + \mathbb{E}^{\mathbb{Q}_f} \left[\sum_{0 \leq s \leq t} e^{-\varrho s} (\Delta f(X)(s) + r \Delta R(s)) \right]. \end{aligned} \quad (4.11)$$

Since $\Delta X(s) = \Delta Y(s) - \Delta R(s)$, we get that

$$\begin{aligned} \Delta f(X)(s) + r \Delta R(s) &= f(X(s)) - f(X(s) - \Delta Y(s)) \\ &\quad - [f(X(s) + \Delta R(s) - \Delta Y(s)) - f(X(s) - \Delta Y(s))] + r \Delta R(s) \\ &= \int_{X(s) - \Delta Y(s)}^{X(s)} f'(u) du + \int_{X(s) - \Delta Y(s)}^{X(s) + \Delta R(s) - \Delta Y(s)} (r - f'(u)) du \\ &\geq 0. \end{aligned} \quad (4.12)$$

Combining (4.11)–(4.12), we get that

$$\mathbb{E}^{\mathbb{Q}_f} [e^{-\varrho t} f(X(t))] + \mathbb{E}^{\mathbb{Q}_f} \left[\int_0^t e^{-\varrho s} \left(h(X(s)) ds + r dR(s) - \frac{1}{2\varepsilon} \psi_f^2(s) ds \right) \right] \geq f(x).$$

Recalling the definition of the cost function J in (3.23) and noting that the function f is bounded as a continuous function on $[0, b]$, by taking $t \rightarrow \infty$, we get that

$$\sup_{\mathbb{Q} \in \mathcal{Q}(x)} J(x, Y, R, \mathbb{Q}; \varepsilon) \geq J(x, Y, R, \mathbb{Q}_f; \varepsilon) \geq f(x).$$

Since x and (Y, R) are arbitrary, it follows that $V(\cdot; \varepsilon) \geq f(\cdot)$. □

Proof of Proposition 4.2: We split the proof into two cases $x \in [0, \beta]$ and $x \in (\beta, b)$. Fix $x \in [0, \beta]$ and an arbitrary $\mathbb{Q} \in \mathcal{Q}(x)$ with an associated process ψ . Recall the notation $\psi_k = \psi_{k_\beta}$. By (4.6), we get

$$\begin{aligned} & \frac{1}{2} \sigma^2 k_\beta''(X(t)) + (m + \sigma \psi(t)) k_\beta'(X(t)) - \varrho k_\beta(X(t)) + h(X(t)) - \frac{1}{2\varepsilon} (\psi(t))^2 \\ & \leq \frac{1}{2} \sigma^2 k_\beta''(X(t)) + (m + \sigma \psi_k(t)) k_\beta'(X(t)) - \varrho k_\beta(X(t)) + h(X(t)) - \frac{1}{2\varepsilon} (\psi_k(t))^2 \\ & = \frac{1}{2} \sigma^2 (k_\beta''(x) + H(x, k_\beta(x), k_\beta'(x))) = 0. \end{aligned} \quad (4.13)$$

Equations (4.9) and (4.10) are given for general admissible strategies $(Y, R) \in \mathcal{A}(x)$ and $\mathbb{Q} \in \mathcal{Q}(x)$. Thus, they hold here as well. Notice that since $x \in [0, \beta]$ and (Y_β, R_β) is a β -reflecting strategy, the processes Y_β and R_β have no jumps and so, $Y_\beta^c = Y_\beta$ and $R_\beta^c = R_\beta$. From (4.9), (4.10), and (4.13), one has,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} [e^{-\varrho t} k_\beta(X(t))] & \leq k_\beta(x) - \mathbb{E}^{\mathbb{Q}} \left[\int_0^t e^{-\varrho s} \left(h(X(s)) ds - \frac{1}{2\varepsilon} \psi^2(s) ds \right) ds \right] \\ & \quad + \mathbb{E}^{\mathbb{Q}} \left[\int_0^t e^{-\varrho s} k_\beta'(X(s)) (dY_\beta(s) - dR_\beta(s)) \right]. \end{aligned} \quad (4.14)$$

Using now the equalities $k_\beta'(0) = 0$ and $k_\beta'(\beta) = r$, driven from (4.7), we get

$$\mathbb{E}^{\mathbb{Q}} [e^{-\varrho t} k_\beta(X(t))] \leq k_\beta(x) - \mathbb{E}^{\mathbb{Q}} \left[\int_0^t e^{-\varrho s} \left(h(X(s)) ds + r dR_\beta(s) - \frac{1}{2\varepsilon} \psi^2(s) ds \right) \right]. \quad (4.15)$$

Recalling the definition of J and noting that the function k_β is bounded as a continuous function on $[0, b]$, by taking $t \rightarrow \infty$, we get that

$$k_\beta(x) \geq \sup_{\mathbb{Q} \in \mathcal{Q}(x)} J(x, Y_\beta, R_\beta, \mathbb{Q}; \varepsilon). \quad (4.16)$$

Notice that all the inequalities in (4.13)–(4.15) hold with equality in case that $\psi = \psi_k$ and $\mathbb{Q} = \mathbb{Q}_k$. Then, together with (4.16), we get that (4.8) holds.

Consider now the case that $\beta < b$ and $x \in (\beta, b]$. From (4.7),

$$k_\beta(x) = r(x - \beta) + k_\beta(\beta).$$

Since the strategy (Y_β, R_β) starts with an instantaneous rejection of $x - \beta_0$, there is an immediate cost of $r(x - \beta)$ and hence for any $\mathbb{Q} \in \mathcal{Q}(x)$,

$$J(x, Y_\beta, R_\beta, \mathbb{Q}; \varepsilon) = r(x - \beta) + J(\beta, Y_\beta, R_\beta, \mathbb{Q}; \varepsilon).$$

Recall that we just showed that

$$k_\beta(\beta) = \sup_{\mathbb{Q} \in \mathcal{Q}(x)} J(\beta, Y_\beta, R_\beta, \mathbb{Q}; \varepsilon) = J(\beta, Y_\beta, R_\beta, \mathbb{Q}_f; \varepsilon).$$

From the last three equalities we get that (4.8) holds for $x \in (\beta, b]$ as well. \square

Before getting to the proof of Proposition 4.2 we provide an auxiliary parametrized ordinary differential equation in order to solve the free boundary one. Fix a parameter $s \in \mathbb{R}$ and consider the following Cauchy problem

$$\begin{cases} (k^{(s)})''(x) + H_F(x, k^{(s)}(x), (k^{(s)})'(x)) = 0, & x \in [0, b], \\ (k^{(s)})'(0) = 0, & k^{(s)}(0) = s, \end{cases} \quad (4.17)$$

where

$$H_F(x, y, z) := H(x, y, F(z)) \quad (4.18)$$

and F is a $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ function that satisfies the following properties: $F(z) = z$ on $[-r, r]$, $|F| \leq 2r$, and $|F'| \leq 1$ and therefore Lipschitz continuous. For example,

$$F(z) = \begin{cases} -3/2r, & z < -2r, \\ (1/2)r + 2z + z^2/(2r), & -2r \leq z < -r, \\ z, & -r \leq z \leq r, \\ -(1/2)r + 2z - z^2/(2r), & r \leq z < 2r, \\ 3/2r, & 2r \leq z. \end{cases}$$

Because the function F and its derivative are bounded, and since the function h is Lipschitz (see the paragraph below (3.21)), H_F is uniformly Lipschitz. Namely, there is a constant c_L such that for every $(x, y, z), (x', y', z') \in [0, b] \times \mathbb{R} \times \mathbb{R}$, one has

$$|H_F(x, y, z) - H_F(x', y', z')| \leq L(|x - x'| + |y - y'| + |z - z'|). \quad (4.19)$$

From [26, Section 0.3.1], (4.17) admits a unique $\mathcal{C}^2([0, b], \mathbb{R})$ solution.

Set

$$\beta^{(s)} := \inf\{x \in (0, b] : (k^{(s)})'(x) \geq r\} \wedge b, \quad (4.20)$$

where we use the convention that $\inf \emptyset = \infty$. The smoothness of $k^{(s)}$ implies that

$$\text{if } \beta^{(s)} < b \text{ then } (k^{(s)})'(\beta^{(s)}) = r. \quad (4.21)$$

The following lemma provides some continuity properties that serve us in the proof of Proposition 4.3.

Lemma 4.2 *The function $s \mapsto (k^{(s)}, (k^{(s)})', (k^{(s)})'')$ is continuous in the uniform norm topology taken on the interval $[0, b]$. Moreover, the mapping $s \mapsto \beta^{(s)}$ is continuous for every s for which either $\beta^{(s)} = b$ or the following two conditions hold $\beta^{(s)} < b$ and $(k^{(s)})''(\beta^{(s)}) \neq 0$. In these cases we conclude that the mapping $s \mapsto (k^{(s)}(\beta^{(s)}), (k^{(s)})'(\beta^{(s)}), (k^{(s)})''(\beta^{(s)}))$ is also continuous.*

Proof: Fix $s \in \mathbb{R}$. One can easily verify that the conditions of [27, Theorem 23] are satisfied for H_F . Therefore, for any $\delta_1 \in \mathbb{R}$

$$\sup_{x \in [0, b]} |k^{(s+\delta_1)}(x) - k^{(s)}(x)| \leq |\delta_1|, \quad (4.22)$$

and the continuity of $s \mapsto k^{(s)}$ is established.

We now turn to showing that $s \mapsto (k^{(s)})'$ is continuous. From (4.17) it follows that for every $x \in [0, b]$

$$\begin{aligned} (k^{(s)})'(x) &= 0 - \int_0^x H_F(y, k^{(s)}(y), (k^{(s)})'(y)) dy, \\ (k^{(s+\delta_1)})'(x) &= 0 - \int_0^x H_F(y, k^{(s+\delta_1)}(y), (k^{(s+\delta_1)})'(y)) dy. \end{aligned}$$

Set $g_{s, \delta_1}(x) := |(k^{(s+\delta_1)})'(x) - (k^{(s)})'(x)|$, $x \in [0, b]$. From (4.19), there exists a constant $L > 0$ independent of s and δ_1 such that

$$g_{s, \delta_1}(x) \leq L \int_0^x \left(|k^{(s+\delta_1)}(y) - k^{(s)}(y)| + g_{s, \delta_1}(y) \right) dy \leq Lb|\delta_1| + L \int_0^x g_{s, \delta_1}(y) dy,$$

where the second inequality follows by (4.22). Now, Grönwall's inequality implies that $\sup_{x \in [0, b]} g_{s, \delta_1}(x) \leq |\delta_1| Lbe^{Lb}$, and the continuity of $s \mapsto (k^{(s)})'$ is established.

Finally, the continuity of $s \mapsto (k^{(s)})''$ follows by the relation $(k^{(s)})''(x) = -H_F(x, k^{(s)}(x), (k^{(s)})'(x))$, the continuity of $s \mapsto (k^{(s)}, (k^{(s)})')$, and the Lipschitz continuity of H_F stated in (4.19).

The rest of the proof is dedicated to the continuity of the function $s \mapsto \beta^{(s)}$ under the conditions mentioned in the lemma. To this end, we fix $s \in \mathbb{R}$ and show that if $\beta^{(s)} = b$ or if $\beta^{(s)} < b$ and $(k^{(s)})'(\beta^{(s)}) < r$, then

$$\limsup_{\delta \rightarrow 0} \beta^{(s+\delta)} \leq \beta^{(s)} \leq \liminf_{\delta \rightarrow 0} \beta^{(s+\delta)}. \quad (4.23)$$

We start with the first inequality. If $\beta^{(s)} = b$ then it is obvious, since all the $\beta^{(u)}$'s are less or equal to b . If $\beta^{(s)} < b$ and $(k^{(s)})''(\beta^{(s)}) \neq 0$, we necessarily have $(k^{(s)})''(\beta^{(s)}) > 0$. Otherwise, $(k^{(s)})''(\beta^{(s)}) < 0$ and from (4.21), we get that $(k^{(s)})'(\beta^{(s)} - \nu) > r$ for sufficiently small $\nu > 0$, a contradiction to the definition of $\beta^{(s)}$.

Using now $(k^{(s)})''(\beta^{(s)}) > 0$, we get that for sufficiently small $\nu > 0$, $(k^{(s)})'(\beta^{(s)} + \nu) > r$. By the continuity of $s \mapsto (k^{(s)})'$, we get that for every δ_2 with sufficiently small absolute value, one has $(k^{(s+\delta_2)})'(\beta^{(s)} + \nu) > r$. Therefore, $\beta^{(s+\delta_2)} < \beta^{(s)} + \nu$ and $\limsup_{\delta \rightarrow 0} \beta^{(s+\delta)} \leq \beta^{(s)} + \nu$. Since $\nu > 0$ can be arbitrary small we get the first inequality on (4.23).

We now turn to proving the second inequality in (4.23). Set $\gamma_1 > 0$ and $\hat{\beta}^{(s)} := \liminf_{\delta \rightarrow 0} \beta^{(s+\delta)}$. Consider a sequence $\{\delta_j\}_j \rightarrow 0$ such that $\{\beta^{(s+\delta_j)}\}_j \rightarrow \hat{\beta}^{(s)}$ and $\beta^{(s+\delta_j)} < b$ for every j . If such

a subsequence does not exist it means that for every δ with sufficiently small absolute value, $\beta^{(s+\delta)} = b$ and the claim is trivial. Since $k^{(s)} \in \mathcal{C}^2([0, b], \mathbb{R})$, we get that for sufficiently large j ,

$$\left| (k^{(s)})'(\beta^{(s+\delta_j)}) - (k^{(s)})'(\hat{\beta}^{(s)}) \right| < \gamma_1. \quad (4.24)$$

Together with the continuity of $s \mapsto (k^{(s)})'$, we get that for sufficiently large j , one has

$$\left| (k^{(s+\delta_j)})'(\beta^{(s+\delta_j)}) - (k^{(s)})'(\beta^{(s+\delta_j)}) \right| < \gamma_1. \quad (4.25)$$

Recall that $\beta^{(s+\delta_j)} < b$. Thus, $(k^{(s+\delta_j)})'(\beta^{(s+\delta_j)}) = r$. From (4.24)–(4.25) we get that

$$\left| (k^{(s)})'(\hat{\beta}^{(s)}) - r \right| < 2\gamma_1.$$

Since $\gamma_1 > 0$ can be arbitrary small, we get that $(k^{(s)})'(\hat{\beta}^{(s)}) = r$ and therefore, $\beta^{(s)} \leq \hat{\beta}^{(s)}$. \square

Proof of Proposition 4.3: Fix $\varepsilon \in (0, \infty)$. We start the proof by arguing uniqueness of the solution and then we move on to showing that a solution indeed exists.

Uniqueness: We first argue that there cannot be two different $\mathcal{C}^2([0, b], \mathbb{R})$ solutions of $\text{HJB}(\varepsilon)$. Notice that the non-linearity of our differential equation prevents us from using the same proof given in [6, Proposition 2.2]. Arguing by contradiction, assume that there are two $\mathcal{C}^2([0, b], \mathbb{R})$ solutions $f_1 \neq f_2$. Without loss of generality, assume that there exists $x \in [0, b]$ such that $f_1(x) > f_2(x)$. Take $y_0 \in \arg \max_{x \in [0, b]} \{f_1(x) - f_2(x)\}$. Clearly, $f_1(y_0) > f_2(y_0)$. Also, $f_1'(0) = f_2'(0) = 0$ and $f_1'(b) = f_2'(b) = r$ since they are both solutions to $\text{HJB}(\varepsilon)$. Now, if $y_0 \in (0, b)$, then first order condition for $f_1 - f_2$ implies that $f_1'(y_0) = f_2'(y_0)$, and therefore it holds whether y_0 is an internal point or an endpoint of $[0, b]$.

By the structure of $\text{HJB}(\varepsilon)$ applied to f_1 , we get

$$f_1''(y_0) + H(y_0, f_1(y_0), f_1'(y_0)) \geq 0. \quad (4.26)$$

The same inequality holds for f_2 . Now, if

$$f_2''(y_0) + H(y_0, f_2(y_0), f_2'(y_0)) = 0, \quad (4.27)$$

then by subtracting (4.27) from (4.26), and recalling that $f_1(y_0) > f_2(y_0)$ and $f_1'(y_0) = f_2'(y_0)$, we get that $f_1''(y_0) > f_2''(y_0)$. Therefore, in a small neighborhood of y_0 (in case that $y_0 = 0$ or $y_0 = b$, we take right- and left-neighborhoods, respectively), $f_1(x) - f_2(x) > f_1(y_0) - f_2(y_0)$, which contradicts the definition of y_0 .

Consider now the case

$$f_2''(y_0) + H(y_0, f_2(y_0), f_2'(y_0)) > 0.$$

Since f_2 is a $\mathcal{C}^2([0, b], \mathbb{R})$ solution of $\text{HJB}(\varepsilon)$, it follows that $0 \leq y_1 < y_0 < y_2 \leq b$, where

$$\begin{aligned} y_1 &:= \sup \{x \in [0, y_0] : f_2''(x) + H(x, f_2(x), f_2'(x)) = 0\} \vee 0, \\ y_2 &:= \inf \{x \in [y_0, b] : f_2''(x) + H(x, f_2(x), f_2'(x)) = 0\} \wedge b. \end{aligned}$$

We use the convention that $\sup \emptyset = -\inf \emptyset = -\infty$. Notice that on the interval (y_1, y_2) , which includes y_0 , $f_2''(x) + H(x, f_2(x), f_2'(x)) > 0$. Since f_2 solves $\text{HJB}(\varepsilon)$, it follows that $f_2' = 0$ on $[y_1, y_2]$ or $f_2' = r$ on $[y_1, y_2]$. In the former case, we get $f_2(y_0) = f_2(y_2)$ and since $f_2'(b) = r$, $y_2 < b$. Since $f_1' \geq 0$ as a solution of $\text{HJB}(\varepsilon)$, we get that $f_1(y_2) - f_2(y_2) \geq f_1(y_0) - f_2(y_0)$, which by the definition of y_0 is in fact an equality. Therefore, $y_2 \in \arg \max_{x \in [0, b]} \{f_1(x) - f_2(x)\}$. Since $y_2 < b$ it follows by the smoothness of f_2 and the definition of y_2 that (4.26) and (4.27) hold with y_2 replacing y_0 . Repeating the same argument proceeding (4.27), with y_2 replacing y_0 yields a contradiction. Now if, $f_2' = r$ on $[y_1, y_2]$, a similar argument hold by using $f_1' \leq r$ and the point y_1 instead of y_2 .

Existence: We now construct a $\mathcal{C}^2([0, b], \mathbb{R})$ solution of (4.7) with some $\beta_\varepsilon \in (0, b]$ that also solves $\text{HJB}(\varepsilon)$. The solution is shown to satisfy also $k'_{\beta_\varepsilon} \leq r$. Repeating the same arguments given in [18, Proposition 6.1], we get that $k''_{\beta_\varepsilon}(x) + H(x, k_{\beta_\varepsilon}(x), k'_{\beta_\varepsilon}(x)) \geq 0$ on $[\beta, b]$ and therefore, k_{β_ε} satisfies $\text{HJB}(\varepsilon)$ and we are done.

To solve (4.7), we hereby consider the Cauchy problem given in (4.17). Since eventually we find a solution to (4.17) that satisfies $0 \leq (k^{(s)})' \leq r$, it also solves the same ordinary differential equation from (4.17) with H replacing H_F . The rest of the proof is performed in two steps. First, we prove the existence of $s^* \in \mathbb{R}$ for which the parameter $\beta^{(s^*)} \in (0, b]$ from (4.20) satisfies the following

$$(k^{(s^*)})'(x) < r \quad \text{on } [0, \beta^{(s^*)}), \quad (4.28)$$

$$(k^{(s^*)})'(\beta^{(s^*)}) = r, \quad (4.29)$$

and in case that $\beta^{(s^*)} < b$, also

$$(k^{(s^*)})''(\beta^{(s^*)}) = 0. \quad (4.30)$$

In the second step we show that $(k^{(s)})' \geq 0$ on $[0, \beta^{(s^*)}]$. Therefore, the function

$$k(x) = \begin{cases} k^{(s^*)}(x), & 0 \leq x < \beta^{(s^*)}, \\ k^{(s^*)}(\beta^{(s^*)}) + r(x - \beta^{(s^*)}), & \beta^{(s^*)} \leq x \leq b, \end{cases}$$

satisfies (4.7), and the proof is done by setting

$$\beta_\varepsilon := \beta^{(s^*)} \quad \text{and} \quad k_{\beta_\varepsilon} := k^{(s^*)}. \quad (4.31)$$

As a conclusion, we get that

$$s^* = k_{\beta_\varepsilon}(0) = V(0; \varepsilon). \quad (4.32)$$

Step 1: Notice that (4.28) holds trivially for every s^* by the definition of $\beta^{(s)}$. Therefore, we only need to check that (4.29) and (4.30) hold.

The first observation is that for sufficiently large s , $\beta^{(s)} < b$ and $(k^{(s)})''(\beta^{(s)}) > 0$. Also, for sufficiently small s , $\beta^{(s_0)} = b$ and $(k^{(s_0)})'(b) < r$. We prove the first part, the second one follows by the same lines and is therefore omitted. Fix $s_1 > (2|m|r + 2\sigma^2\varepsilon r^2 + h(b) + r\sigma^2/b)/\varrho$. Now, the

function $k^{(s_1)}$ is nondecreasing on $[0, b]$. Indeed, arguing by contradiction, assume that there is $x \in (0, b)$ such that $(k^{(s_1)})'(x) < 0$, then consider $y_3 := \inf\{x \in [0, b] : (k^{(s_1)})'(x) < 0\}$. The smoothness of $k^{(s_1)}$ on $[0, b]$ and the initial condition $(k^{(s_1)})'(0) = 0$ imply that $(k^{(s_1)})'(y_3) = 0$. So,

$$\frac{1}{2}\sigma^2(k^{(s_1)})''(y_3) = \varrho k^{(s_1)}(y_3) - h(y_3) \geq \varrho k^{(s_1)}(0) - h(b) = \varrho s_1 - h(b) > 0. \quad (4.33)$$

The first inequality follows since $k^{(s_1)}$ is nondecreasing on $[0, y_3]$. This is a contradiction to the definition of y_3 , since for every sufficiently small $\delta > 0$ we have $(k^{(s_1)})'(y_3 + \delta) > 0$. Therefore, $(k^{(s_1)})' \geq 0$ on $[0, b]$. Combining it with (4.17), $|F| \leq 2r$, $k^{(s_1)}(0) = s_1$, the choice of s_1 , and that h is increasing (see the paragraph below (3.21)), we get that for every $x \in [0, b]$,

$$\begin{aligned} (k^{(s_1)})''(x) &= -\frac{2}{\sigma^2} \left(mF((k^{(s_1)})'(x)) + \frac{1}{2}\sigma^2\varepsilon F^2((k^{(s_1)})'(x)) - \varrho k^{(s_1)}(x) + h(x) \right) \\ &> \frac{2}{\sigma^2} \left(\varrho k^{(s_1)}(0) - 2|m|r - \frac{1}{2}\sigma^2 r^2 \varepsilon - h(b) \right) = \frac{2r}{b}. \end{aligned}$$

In particular $(k^{(s_1)})''(\beta^{(s_1)}) > 0$. Also $(k^{(s_1)})'(b/2) = \int_0^{b/2} (k^{(s_1)})''(u) du \geq r$ and therefore, $\beta^{(s_1)} \leq b/2 < b$.

By the choices of s_0 and s_1 , the following infimum is attained. Set

$$s^* := \inf\{s \in (-s_0, s_1) : \forall u \in (s, s_1), \beta^{(u)} < b \text{ and } (k^{(u)})''(\beta^{(u)}) > 0\}.$$

In case that $\beta^{(s^*)} = b$, then by the definition of s^* , for every $u \in (s^*, s_1)$, $\beta^{(u)} < b$ and therefore, from (4.21) we have $(k^{(u)})'(\beta^{(u)}) = r$. Now, the continuity of the mapping $s \mapsto (k^{(s)})'(\beta^{(s)})$ at $s = s^*$ in this case, provided in Lemma 4.2 implies that $(k^{(s^*)})'(\beta^{(s^*)}) = r$ and (4.29) holds in this case.

In case that $\beta^{(s^*)} < b$ then (4.21) implies that (4.29) holds. We now claim that (4.30) holds, that is, $(k^{(s^*)})''(\beta^{(s^*)}) = 0$. Otherwise, if $(k^{(s^*)})''(\beta^{(s^*)}) < 0$, then by the continuity of the mapping $s \mapsto (\beta^{(s)}, (k^{(s)})''(\beta^{(s)}))$ in this case (see Lemma 4.2), we get that for every sufficiently small $\nu > 0$, $\beta^{(s^*+\nu)} < b$ and $(k^{(s^*+\nu)})''(\beta^{(s^*+\nu)}) < 0$, which contradicts the definition of s^* . If however, $(k^{(s^*)})''(\beta^{(s^*)}) > 0$ then Lemma 4.2 again implies that for every sufficiently small $\nu > 0$, one has $\beta^{(s^*-\nu)} < b$ and $(k^{(s^*-\nu)})''(\beta^{(s^*-\nu)}) > 0$, in contradiction to the definition of s^* . Therefore, $(k^{(s^*)})''(\beta^{(s^*)}) = 0$ and (4.30) holds.

Step 2: In this step we show that

$$(k^{(s^*)})'(x) \geq 0, \quad x \in [0, \beta^{(s^*)}]. \quad (4.34)$$

Arguing by contradiction, assume that there is $y_5 \in (0, \beta^{(s^*)})$ with $(k^{(s^*)})'(y_5) < 0$. We omitted the endpoints where the derivatives are 0 and r . The smoothness of the function $k^{(s^*)}$ together with $(k^{(s^*)})'(0) = 0$ and $(k^{(s^*)})'(\beta^{(s^*)}) = r$ (from the previous step) imply that the following supremum and infimum are attained

$$\begin{aligned} y_4 &:= \inf\{x \in [0, y_5] : \forall y \in (x, y_5), (k^{(s^*)})'(y) \leq 0\}, \\ y_6 &:= \sup\{x \in (y_5, \beta^{(s^*)}) : \forall y \in (x, y_5), (k^{(s^*)})'(y) \leq 0\}. \end{aligned}$$

Also, $(k^{(s^*)})'(y_4) = (k^{(s^*)})'(y_6) = 0$ and $(k^{(s^*)})''(y_4) \leq 0 \leq (k^{(s^*)})''(y_6)$. Substituting these relations in (4.17), we obtain,

$$\varrho k^{(s^*)}(y_4) - h(y_4) \leq 0 \leq \varrho k^{(s^*)}(y_6) - h(y_6). \quad (4.35)$$

Since $(k^{(s^*)})'(x) \leq 0$ on $[y_4, y_6]$, and by the smoothness of $k^{(s^*)}$, there is a subinterval on which $(k^{(s^*)})'(x) < 0$, one has $k^{(s^*)}(y_4) > k^{(s^*)}(y_6)$. Recall that the function h is increasing (see the paragraph below (3.21)) and therefore, $h(y_4) < h(y_6)$. The last two inequalities contradict (4.35) and therefore, (4.34) holds. \square

5 Optimal strategies and equilibria in the games

In Theorem 4.1 we claimed that the minimizer has an optimal strategy, which is a β_ε -reflecting strategy with β_ε given in (4.4). However, we did not argue uniqueness in the sense that for every $\beta \neq \beta_\varepsilon$, the β -reflecting strategy is strictly sub-optimal. Neither did Harrison and Taksar nor Atar and Shifrin. Hence, in the following discussion, which evolves around this issue, we allow $\varepsilon = 0$ and consider $\varepsilon \in [0, \infty)$.

Recall the definition of β_ε from (4.4). Define now,

$$\hat{\beta}_\varepsilon := \sup \{x \in (0, b] : \forall y \leq x, V''(y; \varepsilon) + H(y, V(y; \varepsilon), V'(y; \varepsilon)) = 0\}. \quad (5.1)$$

From (4.5) it follows that $\hat{\beta}_\varepsilon \geq \beta_\varepsilon$. Since on the interval $[\beta_\varepsilon, \hat{\beta}_\varepsilon]$ both of the conditions

$$V''(x; \varepsilon) + H(x, V(x; \varepsilon), V'(x; \varepsilon)) = 0 \quad \text{and} \quad V'(x; \varepsilon) = r \quad (5.2)$$

hold, it follows that V solves (4.7) for any $\beta \in [\beta_\varepsilon, \hat{\beta}_\varepsilon]$. Proposition 4.2 implies that every such β -reflecting strategy is optimal. Thus, the non-uniqueness of the optimal reflecting strategy is equivalent to the existence of non-degenerate interval $[\beta_\varepsilon, \hat{\beta}_\varepsilon]$, on which the equations in (5.1) hold. Combining both of them, we get that $V(x; \varepsilon) = (mr + \sigma^2 r^2 \varepsilon / 2 + h(x)) / \varrho$. Using again $V'(x; \varepsilon) = r$, we get that $h'(x) = \varrho r$. Recall that h is piecewise linear with slopes in the set $\{h_1 \mu_1, \dots, h_I \mu_I\}$ (see the paragraph below (3.21)). Thus, if ϱr does not belong to this set, we get uniqueness of the optimal reflecting strategy. Another sufficient condition would be that m is sufficiently negative such that $(mr + \sigma^2 r^2 \varepsilon / 2 + h(b)) / \varrho \leq 0$, since the value function cannot be non-positive. The arguments above are summarized in the following proposition.

Proposition 5.1 *Fix $\varepsilon \in [0, \infty)$. A β -reflecting strategy is optimal if and only if $\beta \in [\beta_\varepsilon, \hat{\beta}_\varepsilon]$. If*

$$\text{for every } i \in \mathcal{I}, h_i \mu_i \neq \varrho r \quad \text{or} \quad mr + \frac{1}{2} \sigma^2 \varepsilon r^2 + h(b) \leq 0, \quad (5.3)$$

then $\beta_\varepsilon = \hat{\beta}_\varepsilon$ and there is a unique optimal reflecting strategy.

Although there might be weaker assumptions under which uniqueness hold, we do not aim in finding them and contented with the not so restrictive conditions given in (5.3).

We already know that the minimizer has an optimal reflecting strategy. We now show that for every $\epsilon \in (0, \infty)^I$ and ε given through (3.15), both the RSDG and the MSDG have

equilibria. Specifically, each player has an optimal strategy that is good against any strategy of the other player. Recall that Proposition 3.1 connects between the costs in the games. Fix $\hat{x}_0 \in \mathcal{X}$, $\epsilon \in (0, \infty)^I$, and an I -dimensional standard Brownian motion \hat{B} . Set $B = \sigma^{-1} \theta \hat{\sigma} \cdot \hat{B}$ and recall (3.14) and (3.15). Let $(Y_{\bar{\beta}_\epsilon}, R_{\bar{\beta}_\epsilon})$ be an optimal reflecting strategy for the minimizer. Also, let $\mathbb{Q}_V = \mathbb{Q}_{V(\cdot; \epsilon)}$ be the measure driven by $\psi_V = \psi_{V(\cdot; \epsilon)}$ (see (4.6)). Recall the definition of the function γ from (3.26) and the relations given in (3.29)–(3.30). Set the strategy $(\hat{Y}_{\bar{\beta}_\epsilon}, \hat{R}_{\bar{\beta}_\epsilon})$ by

$$\hat{R}_{\bar{\beta}_\epsilon}(t) := R_{\bar{\beta}_\epsilon}(t) \mu_{i^*} e_{i^*} \quad \text{and} \quad \hat{Y}_{\bar{\beta}_\epsilon}(t) := \hat{X}_{\bar{\beta}_\epsilon}(t) - \hat{x}_0 - \hat{m}t - \hat{\sigma} \hat{B}(t) + \hat{R}_{\bar{\beta}_\epsilon}(t), \quad t \in \mathbb{R}_+,$$

where i^* is given in (3.22) and for any $t \in \mathbb{R}_+$,

$$\begin{aligned} \hat{X}_{\bar{\beta}_\epsilon}(t) &:= \gamma(X_{\bar{\beta}_\epsilon}(t)), \\ X_{\bar{\beta}_\epsilon}(t) &= x_0 + mt + \sigma B(t) + Y_{\bar{\beta}_\epsilon}(t) - R_{\bar{\beta}_\epsilon}(t). \end{aligned}$$

Moreover, let $\hat{\mathbb{Q}}_V$ be the measure associated with $\hat{\psi}_V(t) = (\hat{\psi}_{V,1}(t), \dots, \hat{\psi}_{V,I}(t))$, given by

$$\hat{\psi}_{V,i}(t) := \frac{\sigma \psi_V(t) (\theta \hat{\sigma})_i \epsilon_i}{\sum_{j=1}^I (\theta \hat{\sigma})_j^2 \epsilon_j}, \quad i \in \mathcal{I}, \quad t \in \mathbb{R}_+. \quad (5.4)$$

We now claim that the mentioned strategies form equilibria in the games. Using the $\bar{\beta}_\epsilon$ -reflecting strategy in the RSDG it follows that the associated workload dynamics $X_{\bar{\beta}_\epsilon}$ is a reflected diffusion that moves continuously along the interval $[0, \bar{\beta}_\epsilon]$. The function γ continuously maps the workload process to the bold curve from Figure 1. If for example $m = \bar{\beta}_\epsilon$ in figure 1, then the process $\hat{X}_{\bar{\beta}_\epsilon}$ moves continuously along the bold curve between the points A and M , where it is reflected at these two points.

The next theorem generalizes Corollary 3.1 and Theorem 4.1 and provides equilibria in the games.

Theorem 5.1 *Fix $\epsilon \in (0, \infty)$. Using the notation above, the triplets $(Y_{\bar{\beta}_\epsilon}, R_{\bar{\beta}_\epsilon}, \mathbb{Q}_V)$ and $(\hat{Y}_{\bar{\beta}_\epsilon}, \hat{R}_{\bar{\beta}_\epsilon}, \hat{\mathbb{Q}}_V)$ form equilibria in the RSDG and the MSDG, respectively. That is,*

$$\begin{aligned} V(x_0; \epsilon) &= \sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} J(x_0, Y_{\bar{\beta}_\epsilon}, R_{\bar{\beta}_\epsilon}, \mathbb{Q}; \epsilon) = J(x_0, Y_{\bar{\beta}_\epsilon}, R_{\bar{\beta}_\epsilon}, \mathbb{Q}_V; \epsilon) = \inf_{(Y, R) \in \mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{Q}_V; \epsilon) \\ &= \hat{V}(\hat{x}_0; \epsilon) = \sup_{\hat{\mathbb{Q}} \in \hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(x_0, \hat{Y}_{\bar{\beta}_\epsilon}, \hat{R}_{\bar{\beta}_\epsilon}, \hat{\mathbb{Q}}; \epsilon) = \hat{J}(\hat{x}_0, \hat{Y}_{\bar{\beta}_\epsilon}, \hat{R}_{\bar{\beta}_\epsilon}, \hat{\mathbb{Q}}_V; \epsilon) = \inf_{(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_V; \epsilon). \end{aligned}$$

Remark 5.1 *Notice that the optimal policies for the minimizer in the MSDG has the same structure as the optimal policy in the Brownian control problem given in [6]. The only difference emerge from the cutoff point β_ϵ , which affects the point of reflection on the curve γ . Such a result is not obvious due to the non-stationarity structure of the problem caused by the existence of the maximizer player. Furthermore, the structures of the equilibria share similarities with the optimal policies in the differential game given in [2]. More specifically, [2, Theorem 3.3] states that the minimizer's optimal strategy in the one-dimensional deterministic differential game is a reflecting strategy (called there a 'barrier strategy'). Under optimality, the maximal player in the same game uses a drift change that is illustrated more compactly in [3, Section 3, (30)–(31)]. The relationship between the multidimensional and the one-dimensional games is given in [2, Appendix A].*

Proof of Theorem 5.1: In this proof we do not check admissibility. It can be verified easily the same way as in Proposition 3.1 using Corollary 4.1. The first two equalities follow from the optimality of $(Y_{\bar{\beta}_\varepsilon}, R_{\bar{\beta}_\varepsilon})$ and from (4.8). The third equality follows since on the one hand, by the definition of V ,

$$V(x_0; \varepsilon) \geq \inf_{(Y, R) \in \mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{Q}_V; \varepsilon)$$

and on the other hand, the reversed inequality follows from Proposition 4.1 since by Theorem 4.1, $V \in \mathcal{C}^2([0, b], \mathbb{R})$. The forth equality follows by Corollary 3.1. For the fifth and sixth equalities notice that $\hat{V}(\hat{x}_0; \varepsilon) = V(x_0; \varepsilon)$ and that

$$\hat{V}(\hat{x}_0; \varepsilon) \leq \sup_{\hat{\mathbb{Q}} \in \hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}_{\bar{\beta}_\varepsilon}, \hat{R}_{\bar{\beta}_\varepsilon}, \hat{\mathbb{Q}}; \varepsilon) \leq \sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} J(x_0, Y_{\bar{\beta}_\varepsilon}, R_{\bar{\beta}_\varepsilon}, \mathbb{Q}; \varepsilon) = V(x_0; \varepsilon),$$

where the second inequality follows by (3.32). For the last equality, notice that we already established

$$\hat{V}(\hat{x}_0; \varepsilon) = \sup_{\hat{\mathbb{Q}} \in \hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(x_0, \hat{Y}_{\bar{\beta}_\varepsilon}, \hat{R}_{\bar{\beta}_\varepsilon}, \hat{\mathbb{Q}}; \varepsilon) \geq \hat{J}(\hat{x}_0, \hat{Y}_{\bar{\beta}_\varepsilon}, \hat{R}_{\bar{\beta}_\varepsilon}, \hat{\mathbb{Q}}_V; \varepsilon) \geq \inf_{(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_V; \varepsilon).$$

Therefore, it is sufficient to show that $\inf_{(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_V; \varepsilon) \geq V(x_0; \varepsilon)$, which holds by (3.28) (replace the subindex $*$ with V) as follows,

$$\inf_{(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_V; \varepsilon) \geq \inf_{(Y, R) \in \mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{Q}_V; \varepsilon) = V(x_0; \varepsilon).$$

□

6 Dependency on the ambiguity parameters

In this section we study the dependence of the value functions and the optimal cutoffs on the ambiguity parameters. We show continuity and that as $\varepsilon \rightarrow 0$ our model converges to the risk-neutral model, studied by Harrison and Taksar [18] and by Atar and Shifrin [6]. For this, recall the definition of $V(\cdot; 0)$ given in Remark 3.2.

Theorem 6.1 *The mapping $[0, \infty) \ni \varepsilon \mapsto (V(\cdot; \varepsilon), V'(\cdot; \varepsilon))$ is increasing and continuous in the uniform norm topology taken on the interval $[0, b]$. Moreover, there is a constant $C > 0$ such that for every $\varepsilon \in (0, \infty)$, $\sup_{x \in [0, b]} |V(x; \varepsilon) - V(x; 0)| \leq C\varepsilon$. Also, $\lim_{\varepsilon \rightarrow \infty} V(\cdot; \varepsilon) = \infty$, uniformly on $[0, b]$. Finally, consider the relations \prec and \preceq on $(0, \infty)^I \times (0, \infty)^I$ given by,*

$$\epsilon \prec \epsilon' \quad (\text{resp., } \preceq) \quad \text{if and only if} \quad \sum_{i=1}^I (\theta \hat{\sigma})_i^2 \epsilon_i < \sum_{i=1}^I (\theta \hat{\sigma})_i^2 \epsilon'_i \quad (\text{resp., } \preceq).$$

Then the mapping $(0, \infty)^I \ni \epsilon \mapsto \hat{V}(\cdot; \epsilon)$ is increasing w.r.t. \prec and continuous in the uniform norm topology taken on \mathcal{X} .

Proof: The last part of the theorem merely follows by the first one, (3.15), and Corollary 3.1. Therefore it is omitted and we turn to proving the first part of the theorem. We start by showing the monotonicity and continuity of the mapping $(0, \infty) \ni \varepsilon \mapsto V(\cdot; \varepsilon)$. The proof for $\varepsilon = 0$ is given separately. Fix $0 < \varepsilon_2 < \varepsilon_1$. Denote by $\mathbb{Q}_i = \mathbb{Q}_{V(\cdot; \varepsilon_i)}$ and $\psi_i = \psi_{V(\cdot; \varepsilon_i)}$, $i = 1, 2$. Then, for every $x_0 \in [0, b]$ one has

$$\begin{aligned}
V(x_0; \varepsilon_1) &= \inf_{(Y, R) \in \mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{Q}_1; \varepsilon_1) \\
&= \inf_{(Y, R) \in \mathcal{A}(x_0)} \left[J(x_0, Y, R, \mathbb{Q}_1; \varepsilon_2) + \frac{1}{2} \left(\frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1} \right) \int_0^\infty e^{-\varrho t} \mathbb{E}^{\mathbb{Q}_1} [\psi_1^2(t)] dt \right] \\
&\leq \inf_{(Y, R) \in \mathcal{A}(x_0)} \left[J(x_0, Y, R, \mathbb{Q}_1; \varepsilon_2) + \frac{\varepsilon_1 \sigma^2 r^2}{2\varepsilon_2 \varrho} (\varepsilon_1 - \varepsilon_2) \right] \\
&\leq \inf_{(Y, R) \in \mathcal{A}(x_0)} \sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} [J(x_0, Y, R, \mathbb{Q}; \varepsilon_2)] + \frac{\varepsilon_1 \sigma^2 r^2}{2\varepsilon_2 \varrho} (\varepsilon_1 - \varepsilon_2) \\
&= V(x_0; \varepsilon_2) + \frac{\varepsilon_1 \sigma^2 r^2}{2\varepsilon_2 \varrho} (\varepsilon_1 - \varepsilon_2).
\end{aligned}$$

The first equality follows by Theorem 5.1. The second equality follows by (3.23). The first inequality follows since by (4.6), $\psi_1(t) = \varepsilon \sigma V'(X(t); \varepsilon_1)$ and since $V'(x; \varepsilon_1) \leq r$, see Corollary 4.1. The second inequality is trivial and finally, the last equality follows by the definition of V , see (3.24).

On the other hand,

$$\begin{aligned}
V(x_0; \varepsilon_1) &\geq \inf_{(Y, R) \in \mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{Q}_2; \varepsilon_1) \\
&= \inf_{(Y, R) \in \mathcal{A}(x_0)} \left[J(x_0, Y, R, \mathbb{Q}_2; \varepsilon_2) + \frac{1}{2} \left(\frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1} \right) \int_0^\infty e^{-\varrho t} \mathbb{E}^{\mathbb{Q}_2} [\psi_2^2(t)] dt \right] \\
&> \inf_{(Y, R) \in \mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{Q}_2; \varepsilon_2) \\
&= V(x_0; \varepsilon_2).
\end{aligned}$$

The first inequality follows by the definition of V . The first equality follows by (3.23). The strict inequality follows since $\varepsilon_1 > \varepsilon_2$ and since by (4.6), $\psi_1(t) = \varepsilon \sigma V'(X(t); \varepsilon_1)$ and since $V'(x; \varepsilon_1) \geq 0$, see Corollary 4.1. Obviously, with \mathbb{Q}_2 -probability zero, $V'(X(t); \varepsilon_1) = 0$ for almost every $t \in \mathbb{R}_+$ w.r.t. Lebesgue measure. The last equality follows by Theorem 5.1. Combining the last two sets of relations, we obtain,

$$V(x_0; \varepsilon_2) < V(x_0; \varepsilon_1) \leq V(x_0; \varepsilon_2) + \frac{\varepsilon_1 \sigma^2 r^2}{2\varepsilon_2 \varrho} (\varepsilon_1 - \varepsilon_2)$$

and the monotonicity and continuity of $\varepsilon \mapsto V(\cdot; \varepsilon)$ is proven on the interval $(0, \infty)$. The

monotonicity at $\varepsilon = 0$ follows by the following sequence of relations,

$$\begin{aligned}
V(x_0; 0) &= \inf_{(Y, R) \in \mathcal{A}(x_0)} \mathbb{E}^{\mathbb{P}} \left[\int_0^\infty e^{-\varrho t} (h(X(t))dt + r dR(t)) \right] \\
&= \inf_{(Y, R) \in \mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{P}; \varepsilon_1) \\
&\leq \inf_{(Y, R) \in \mathcal{A}(x_0)} \sup_{\mathbb{Q} \in \mathcal{Q}(x_0)} J(x_0, Y, R, \mathbb{Q}; \varepsilon_1) \\
&= V(x_0; \varepsilon_1),
\end{aligned} \tag{6.1}$$

where $x_0 \in [0, b]$ and $\varepsilon_1 \in (0, \infty)$. The first equality follows by the definition of the cost in the risk-neutral case, see (3.25). The second equality follows since $L^\varrho(\mathbb{P}|\mathbb{P}) = 0$.

We now turn to proving the continuity at $\varepsilon = 0$. Notice that in the arguments above, the inequality $\varepsilon_2 > 0$ cannot be relaxed to $\varepsilon_2 \geq 0$ since we divide by ε_2 . Therefore, we come up with another proof for continuity at 0. Recall that the arguments in Section 5 included the case where ε is zero. Fix an optimal cutoff in the risk-neutral problem $\bar{\beta}_0 \in [\beta_0, \hat{\beta}_0]$. Fix also $\varepsilon > 0$ and set $\mathbb{Q}^\varepsilon := \mathbb{Q}_{V(\cdot; \varepsilon)}$ and $\psi^\varepsilon := \psi_{V(\cdot; \varepsilon)}$. From (6.1) and Theorem 5.1 one has,

$$V(x_0; 0) \leq V(x_0; \varepsilon) = J(x_0, Y_{\bar{\beta}_0}, R_{\bar{\beta}_0}, \mathbb{Q}^\varepsilon; \varepsilon).$$

Recall that by (3.23) and (4.6),

$$\begin{aligned}
&J(x_0, Y_{\bar{\beta}_0}, R_{\bar{\beta}_0}, \mathbb{Q}^\varepsilon; \varepsilon) \\
&= \mathbb{E}^{\mathbb{Q}^\varepsilon} \left[\int_0^\infty e^{-\varrho t} (h(X_{\bar{\beta}_0}(t))dt + r dR_{\bar{\beta}_0}(t)) \right] - \varepsilon \sigma^2 \int_0^\infty e^{-\varrho t} (V'(X_{\bar{\beta}_0}(s); \varepsilon))^2 ds,
\end{aligned} \tag{6.2}$$

where

$$X_{\bar{\beta}_0}(t) = x_0 + mt + \sigma B(t) + Y_{\bar{\beta}_0}(t) - R_{\bar{\beta}_0}(t), \quad t \in \mathbb{R}_+. \tag{6.3}$$

Since $0 \leq V'(\cdot; \varepsilon) \leq r$, the last term on the r.h.s. of (6.2) goes to zero as $\varepsilon \rightarrow 0$. Hence, in order to show the limit $\lim_{\varepsilon \rightarrow 0} V(x_0; \varepsilon) = V(x_0; 0)$ it is sufficient to show that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}^{\mathbb{Q}^\varepsilon} \left[\int_0^\infty e^{-\varrho t} (h(X_{\bar{\beta}_0}(t))dt + r dR_{\bar{\beta}_0}(t)) \right] = \mathbb{E}^{\mathbb{P}} \left[\int_0^\infty e^{-\varrho t} (h(X_{\bar{\beta}_0}(t))dt + r dR_{\bar{\beta}_0}(t)) \right]. \tag{6.4}$$

The proof of (6.4) relies on the continuity property of the Skorokhod mapping in addition to a coupling argument. We consider a probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \check{\mathbb{P}})$ that supports a one-dimensional standard Brownian motion B , adapted to the filtration $\{\mathcal{F}_t\}$, such that for every $0 \leq s < t$, $B(t) - B(s)$ is independent of \mathcal{F}_s under $\check{\mathbb{P}}$. Recall Definition 4.1 and consider the following processes,

$$\begin{aligned}
(X^\varepsilon, Y^\varepsilon, R^\varepsilon)(t) &:= \Gamma_{[0, \bar{\beta}_0]} \left(x_0 + m \cdot + \int_0^\cdot \varepsilon \sigma^2 V'(X^\varepsilon(s); \varepsilon)(s) ds + \sigma B(\cdot) \right) (t), \\
(X^0, Y^0, R^0)(t) &:= \Gamma_{[0, \bar{\beta}_0]} (x_0 + m \cdot + \sigma B(\cdot)) (t).
\end{aligned}$$

Notice that by (4.6), (X^0, Y^0, R^0) (resp., $(X^\varepsilon, Y^\varepsilon, R^\varepsilon)$) has the same distribution under the measure $\check{\mathbb{P}}$ as $(X_{\bar{\beta}_0}, Y_{\bar{\beta}_0}, R_{\bar{\beta}_0})$ given in (6.3) under the measure \mathbb{P} (resp., \mathbb{Q}^ε , see (3.19)). Hence, (6.4) follows once we show that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}^{\check{\mathbb{P}}} \left[\int_0^\infty e^{-\varrho t} (h(X^\varepsilon(t))dt + r dR^\varepsilon(t)) \right] = \mathbb{E}^{\check{\mathbb{P}}} \left[\int_0^\infty e^{-\varrho t} (h(X^0(t))dt + r dR^0(t)) \right]. \tag{6.5}$$

Now,

$$\begin{aligned}
& \left| \mathbb{E}^{\mathbb{P}} \left[\int_0^\infty e^{-\varrho t} (h(X^\varepsilon(t)) dt + r dR^\varepsilon(t)) \right] - \mathbb{E}^{\mathbb{P}} \left[\int_0^\infty e^{-\varrho t} (h(X^0(t)) dt + r dR^0(t)) \right] \right| \quad (6.6) \\
&= \left| \mathbb{E}^{\mathbb{P}} \left[\int_0^\infty e^{-\varrho t} [h(X^\varepsilon(t)) - h(X^0(t))] dt + \int_0^\infty \varrho r e^{-\varrho t} [R^\varepsilon(t) - R^0(t)] dt + [e^{-\varrho t} (R^\varepsilon(t) - R^0(t))] \Big|_0^\infty \right] \right| \\
&\leq \mathbb{E}^{\mathbb{P}} \left[\int_0^\infty e^{-\varrho t} |h(X^\varepsilon(t)) - h(X^0(t))| dt + \int_0^\infty \varrho r e^{-\varrho t} |R^\varepsilon(t) - R^0(t)| dt \right].
\end{aligned}$$

The last inequality follows since $\lim_{t \rightarrow \infty} e^{-\varrho t} R^\varepsilon(t) = 0$, \mathbb{P} -a.s. and similarly for R^0 . Indeed, consider $w(\cdot) = x_0 + m \cdot + \int_0^\cdot \varepsilon \sigma^2 V'(X^\varepsilon(s); \varepsilon)(s) ds + \sigma B(\cdot)$ and $\tilde{w} = 0$ in Lemma 4.1. Recalling the bound $0 \leq V'(\cdot, \varepsilon) \leq r$, it follows that $R^\varepsilon(t) \leq c_S(x_0 + (m + \varepsilon \sigma^2 r)t + \sigma \sup_{0 \leq s \leq t} |B(s)|)$ and the limit follows since $\lim_{t \rightarrow \infty} e^{-\varrho t} \sup_{0 \leq s \leq t} |B(s)| = 0$, \mathbb{P} -a.s. Using the Lipschitz continuity of h , again $0 \leq V'(\cdot, \varepsilon) \leq r$, and Lemma 4.1 with $w(\cdot) = x_0 + m \cdot + \int_0^\cdot \varepsilon \sigma^2 V'(X^\varepsilon(s); \varepsilon)(s) ds + \sigma B(\cdot)$ and $\tilde{w}(\cdot) = x_0 + m \cdot + \sigma B(\cdot)$, we get that for every $t \in \mathbb{R}_+$,

$$\sup_{s \in [0, t]} \|(X^\varepsilon, Y^\varepsilon, R^\varepsilon)(s) - (X^0, Y^0, R^0)(s)\| \leq \varepsilon c_S \sigma^2 r t.$$

Recalling that h is Lipschitz, we get that the r.h.s. of (6.6) is bounded above by $C\varepsilon(1/\varrho + r)/\varrho$, where the constant $C > 0$ depends on c_S , $\sigma^2 r$, and the Lipschitz constant of h . This implies (6.5). Recalling (6.2) and the estimations above, we get the bound

$$\sup_{x \in [0, b]} |V(x; \varepsilon) - V(x; 0)| \leq (\sigma^2 r^2 + C\varepsilon(1/\varrho + r)/\varrho)\varepsilon$$

as required.

We now turn to proving the limit $\lim_{\varepsilon \rightarrow \infty} V(\cdot; \varepsilon) = \infty$. For this, consider the strategy of the maximizer given by \mathbb{Q}_ε , which is associated with $\psi_\varepsilon(s) = \varepsilon^{1/4}$ for every $s \in \mathbb{R}_+$. Given that and recalling the presentation of the dynamics provided in (3.19), it follows that,

$$X(t) = x_0 + (m + \sigma \varepsilon^{1/4})t + \sigma B^{\mathbb{Q}_\varepsilon}(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+, \quad (6.7)$$

where $B^{\mathbb{Q}_\varepsilon}$ is an $\{\mathcal{F}_t\}$ -one-dimensional standard Brownian motion under \mathbb{Q}_ε . His cost function is given by

$$\begin{aligned}
J(x_0, Y, R, \mathbb{Q}_\varepsilon; \varepsilon) &= \mathbb{E}^{\mathbb{Q}} \left[\int_0^\infty e^{-\varrho t} \left(h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi^2(t) dt \right) \right] \\
&= \mathbb{E}^{\mathbb{Q}_\varepsilon} \left[\int_0^\infty e^{-\varrho t} (h(X(t)) dt + r dR(t)) \right] - \frac{1}{2\varrho^2 \varepsilon^{1/2}}.
\end{aligned}$$

Since the last term is constant, we can think as if the minimizer is facing a risk-neutral problem with the drift mt replaced by $(m + \sigma \varepsilon^{1/4})t$, where his cost function has an additional deterministic term, which vanishes as $\varepsilon \rightarrow \infty$. Therefore, we analyze only the limit of the expectation in the second line of the above. Consider an optimal β -reflecting strategy (Y_β, R_β) , which exists in the risk-neutral problem. From the dynamics in (6.7), we get for every $t \in \mathbb{R}_+$,

$$R_\beta(t) \geq x_0 - X(t) + (m + \sigma \varepsilon^{1/4})t + \sigma B^{\mathbb{Q}_\varepsilon}(t) + Y_\beta(t) \geq -b + (m + \sigma \varepsilon^{1/4})t + \sigma B^{\mathbb{Q}_\varepsilon}(t).$$

The same arguments that lead to (6.6), imply that

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}_\varepsilon} \left[\int_0^\infty e^{-\varrho t} (h(X(t))dt + rdR_\beta(t)) \right] &= \mathbb{E}^{\mathbb{Q}_\varepsilon} \left[\int_0^\infty e^{-\varrho t} (h(X(t)) + \varrho r R_\beta(t))dt \right] \\ &\geq \mathbb{E}^{\mathbb{Q}_\varepsilon} \left[\int_0^\infty e^{-\varrho t} (h(X(t)) + \varrho r (-b + (m + \sigma \varepsilon^{1/4})t + \sigma B^{\mathbb{Q}_\varepsilon}(t)))dt \right].\end{aligned}$$

Standard estimates give that the last term goes to ∞ together with ε .

□

From Theorem 5.1, Corollary 4.1, and (4.6) it is easy to see the continuity of the maximizer's optimal control w.r.t. ε . Also, it is clear that as $\varepsilon \rightarrow 0$, $\psi_{V(\cdot; \varepsilon)} \rightarrow 0$, uniformly on $[0, b]$ and therefore, the maximizer's optimal strategy is relatively close to keep the original measure as is. Since we do not have explicit expressions for the cutoff points of the reflecting strategies β_ε and $\hat{\beta}_\varepsilon$ described in Section 5, the situation with the minimizer's optimal strategy is more subtle and is studied now. The next theorem provides a continuity property of the optimal cutoff points β_ε and $\hat{\beta}_\varepsilon$ w.r.t. ε . Recall the sufficient conditions given in (5.3) for uniqueness of the reflecting strategy.

Theorem 6.2 *For any given $\varepsilon \in [0, \infty)$,*

$$\beta_\varepsilon \leq \liminf_{\delta \rightarrow 0} \beta_{\varepsilon+\delta} \leq \limsup_{\delta \rightarrow 0} \hat{\beta}_{\varepsilon+\delta} \leq \hat{\beta}_\varepsilon. \quad (6.8)$$

Hence, in case that $\hat{\beta}_\varepsilon = \beta_\varepsilon$ (for the given ε), that is, if there is a unique optimal reflecting strategy, one has,

$$\lim_{\delta \rightarrow 0} \beta_{\varepsilon+\delta} = \lim_{\delta \rightarrow 0} \hat{\beta}_{\varepsilon+\delta} = \hat{\beta}_\varepsilon. \quad (6.9)$$

As a preparation for the proof, we present an auxiliary function defined for every $\varepsilon \in [0, \infty)$. Recall the definition of $k^{(s)}$ given in (4.17) and (4.32). Set $l^{(\varepsilon)}(x) = k^{(V(0; \varepsilon))}(x)$, $x \in [0, b]$. More explicitly,

$$\begin{cases} (l^{(\varepsilon)})''(x) + H_F(x, l^{(\varepsilon)}(x), (l^{(\varepsilon)})'(x)) = 0, & x \in [0, b], \\ (l^{(\varepsilon)})'(0) = 0, \quad l^{(\varepsilon)}(0) = V(0; \varepsilon), \end{cases} \quad (6.10)$$

with the same F given after (4.17). The next lemma provides a continuity property of $l^{(\varepsilon)}$ as a function of ε .

Lemma 6.1 *The mapping $[0, \infty) \ni \varepsilon \mapsto (l^{(\varepsilon)}(\cdot), (l^{(\varepsilon)})'(\cdot))$ is continuous in the uniform norm topology taken on the interval $[0, b]$.*

Proof: Fix $\varepsilon_1, \varepsilon_2 \in \mathbb{R}_+$. Set $f_1 = l^{(\varepsilon_1)}$ and $f_2 = l^{(\varepsilon_2)}$. Then,

$$\begin{aligned}f_1'(x) &= - \int_0^x H_F^{\varepsilon_1}(y, f_1(y), f_1'(y))dy, \\ f_2'(x) &= - \int_0^x H_F^{\varepsilon_2}(y, f_2(y), f_2'(y))dy = - \int_0^x H_F^{\varepsilon_1}(y, f_2(y), f_2'(y))dy + \int_0^x (\varepsilon_2 - \varepsilon_1)F(f_2'(y)),\end{aligned}$$

where the index ε_i in $H_F^{\varepsilon_i}$ emphasizes its dependence on ε_i , $i = 1, 2$. From (4.19) it follows that there exists a constant $L > 0$ independent of $\varepsilon_1, \varepsilon_2$ such that

$$|f'_1(x) - f'_2(x)| \leq L \left(\int_0^x [|f_1(y) - f_2(y)| + |f'_1(y) - f'_2(y)|] dy + |\varepsilon_1 - \varepsilon_2| \right).$$

Also, since $f_i(0) = V(0; \varepsilon_i)$,

$$|f_1(x) - f_2(x)| \leq |V(0; \varepsilon_1) - V(0; \varepsilon_2)| + \int_0^x |f'_1(y) - f'_2(y)| dy.$$

From the last two inequalities and Grönwall's inequality we get that there is a constant $C > 0$ independent of $\varepsilon_1, \varepsilon_2$, and x , such that for every $x \in [0, b]$

$$|f_1(x) - f_2(x)| + |f'_1(x) - f'_2(x)| \leq C (|V(0; \varepsilon_1) - V(0; \varepsilon_2)| + |\varepsilon_1 - \varepsilon_2|).$$

Recalling from Theorem 6.1 that $\varepsilon \mapsto V(\cdot; \varepsilon)$ is continuous, we get that $\sup_{0 \leq x \leq b} (|f_1(x) - f_2(x)| + |f'_1(x) - f'_2(x)|)$ is uniformly bounded by a function of $(\varepsilon_1, \varepsilon_2)$ that goes to zero as $(\varepsilon_1 - \varepsilon_2) \rightarrow 0$. □

Proof of Theorem 6.2: For any $\varepsilon' \in [0, \infty)$, the function $V(\cdot; \varepsilon')$ satisfies (4.5) and therefore also (6.10) on $x \in [0, \hat{\beta}_{\varepsilon'}]$. Uniqueness of the solution implies that

$$V(x; \varepsilon') = l^{(\varepsilon')}(x), \quad x \in [0, \hat{\beta}_{\varepsilon'}], \quad (6.11)$$

see [26, Section 0.3.1].

To get the first inequality in (6.8) notice that from (6.11), the definition of β_ε in (4.4), the definition of $\beta^{(s)}$ in (4.20), and recalling that $l^{(\varepsilon)} = k^{(V(0; \varepsilon))}$, it follows that for any $\varepsilon' \in [0, \infty)$, $\beta_{\varepsilon'} = \beta^{(V(0; \varepsilon'))}$. From the continuity of $\varepsilon \mapsto V(\cdot; \varepsilon)$, see Theorem 6.1, and the second inequality in (4.23), we get the first inequality in (6.8).

The second inequality in (6.8) is trivial and follows since $\beta_{\varepsilon+\delta} \leq \hat{\beta}_{\varepsilon+\delta}$. We now turn to proving the last inequality in (6.8). Fix $\varepsilon \in [0, \infty)$. From Theorem 6.1 and Lemma 6.1 we have

$$\lim_{\delta \rightarrow 0} V(\cdot; \varepsilon + \delta) = V(\cdot; \varepsilon) \quad \text{and} \quad \lim_{\delta \rightarrow 0} l^{(\varepsilon+\delta)}(\cdot) = l^{(\varepsilon)}(\cdot),$$

uniformly on $[0, b]$. Together with (6.11) applied to $\varepsilon' = \varepsilon + \delta$, one has $V(x; \varepsilon + \delta) = l^{(\varepsilon+\delta)}(x)$, $x \in [0, \hat{\beta}_{\varepsilon+\delta}]$. Taking $\delta \rightarrow 0$, we get that $V(x; \varepsilon) = l^{(\varepsilon)}(x)$ for $x \in [0, \alpha_\varepsilon)$, where $\alpha_\varepsilon := \limsup_{\delta \rightarrow 0} \hat{\beta}_{\varepsilon+\delta}$. Since both functions are continuous, the equality holds for α_ε as well. By the definition of β_ε , see (5.1), we get that (6.8) holds. □

References

- [1] R. Atar and A. Biswas. Control of the multiclass $G/G/1$ queue in the moderate deviation regime. *Ann. Appl. Probab.*, 24(5):2033–2069, 2014.

- [2] R. Atar and A. Cohen. A differential game for a multiclass queueing model in the moderate-deviation heavy-traffic regime. *Math. Oper. Res.*, 41(4):1354–1380, 2016.
- [3] R. Atar and A. Cohen. Asymptotically optimal control for a multiclass queueing model in the moderate deviation heavy traffic regime. *Ann. Appl. Probab.*, to appear, 2017.
- [4] R. Atar and G. Mendelson. On the non-Markovian multiclass queue under risk-sensitive cost. *Queueing Syst.*, 84(3-4):265–278, 2016.
- [5] R. Atar and S. Saha. The generalized $c\mu$ rule in the moderate deviation regime. *arXiv preprint arXiv:1512.00620*, 2016.
- [6] R. Atar and M. Shifrin. An asymptotic optimality result for the multiclass queue with finite buffers in heavy traffic. *Stochastic Systems*, 4(2):556–603, 2014.
- [7] E. Bayraktar and Y. Zhang. Minimizing the probability of lifetime ruin under ambiguity aversion. *SIAM J. Control Optim.*, 53(1):58–90, 2015.
- [8] S. L. Bell and R. J. Williams. Dynamic scheduling of a system with two parallel servers in heavy traffic with resource pooling: asymptotic optimality of a threshold policy. *Ann. Appl. Probab.*, 11(3):608–649, 2001.
- [9] A. Biswas. Risk-sensitive control for the multiclass many-server queues in the moderate deviation regime. *Math. Oper. Res.*, 39(3):908–929, 2014.
- [10] J. Blanchet, C. Dolan, and H. Lam. Robust rare-event performance analysis with natural non-convex constraints. In *Proceedings of the 2014 Winter Simulation Conference*, pages 595–603. IEEE Press, 2014.
- [11] A. Budhiraja and A. P. Ghosh. Diffusion approximations for controlled stochastic networks: an asymptotic bound for the value function. *Ann. Appl. Probab.*, 16(4):1962–2006, 2006.
- [12] A. Budhiraja and A. P. Ghosh. Controlled stochastic networks in heavy traffic: convergence of value functions. *Ann. Appl. Probab.*, 22(2):734–791, 2012.
- [13] A. Cohen. Asymptotically optimal control for a multiclass M/M/1 queueing problem with model uncertainty. *Work in progress*, 2014.
- [14] L. P. Hansen, T. J. Sargent, G. Turmuhambetova, and N. Williams. Robust control and model misspecification. *J. Econom. Theory*, 128(1):45–90, 2006.
- [15] J. M. Harrison. Brownian models of queueing networks with heterogeneous customer populations. In *Stochastic differential systems, stochastic control theory and applications (Minneapolis, Minn., 1986)*, volume 10 of *IMA Vol. Math. Appl.*, pages 147–186. Springer, New York, 1988.
- [16] J. M. Harrison. Brownian models of open processing networks: canonical representation of workload. *Ann. Appl. Probab.*, 10(1):75–103, 2000.
- [17] J. M. Harrison. A broader view of Brownian networks. *Ann. Appl. Probab.*, 13(3):1119–1150, 2003.
- [18] J. M. Harrison and M. I. Taksar. Instantaneous control of Brownian motion. *Math. Oper. Res.*, 8(3):439–453, 1983.
- [19] J. M. Harrison and J. A. Van Mieghem. Dynamic control of Brownian networks: state space collapse and equivalent workload formulations. *Ann. Appl. Probab.*, 7(3):747–771, 1997.
- [20] J. M. Harrison and R. J. Williams. Workload reduction of a generalized Brownian network. *Ann. Appl. Probab.*, 15(4):2255–2295, 2005.
- [21] A. Jain, A. E. B. Lim, and J. G. Shanthikumar. On the optimality of threshold control in queues with model uncertainty. *Queueing Syst.*, 65(2):157–174, 2010.
- [22] L. Kruk, J. Lehoczy, K. Ramanan, and S. Shreve. An explicit formula for the Skorokhod map on $[0, a]$. *Ann. Probab.*, 35(5):1740–1768, 2007.
- [23] H. Lam. Robust sensitivity analysis for stochastic systems. *Math. Oper. Res.*, 41(4):1248–1275, 2016.
- [24] P. J. Maenhout. Robust portfolio rules and asset pricing. *Review of financial studies*, 17(4):951–983, 2004.
- [25] E. Plambeck, S. Kumar, and J. M. Harrison. A multiclass queue in heavy traffic with throughput time constraints: asymptotically optimal dynamic controls. *Queueing Syst.*, 39(1):23–54, 2001.

- [26] A. D. Polyanin and V. F. Zaitsev. *Handbook of exact solutions for ordinary differential equations*. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2003.
- [27] M. H. Protter and H. F. Weinberger. *Maximum principles in differential equations*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1967.
- [28] M. Sirbu. A note on the strong formulation of stochastic control problems with model uncertainty. *Electron. Commun. Probab.*, 19:no. 81, 10, 2014.
- [29] J. Stoer and R. Bulirsch. *Introduction to numerical analysis*. Springer-Verlag, New York-Heidelberg, 1980. Translated from the German by R. Bartels, W. Gautschi and C. Witzgall.
- [30] D. W. Stroock. *Lectures on stochastic analysis: diffusion theory*, volume 6 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1987.